cryptography

MDPI

Article

Robust Secure Authentication and Data Storage with Perfect Secrecy

Sebastian Baur * and Holger Boche

Institute of Theoretical Information Technology, Technical University of München, 80333 München, Germany; boche@tum.de

* Correspondence: s.j.baur@tum.de

Received: 29 January 2018; Accepted: 6 April 2018; Published: 10 April 2018

Abstract: We consider an authentication process that makes use of biometric data or the output of a physical unclonable function (PUF), respectively, from an information theoretical point of view. We analyse different definitions of achievability for the authentication model. For the secrecy of the key generated for authentication, these definitions differ in their requirements. In the first work on PUF based authentication, weak secrecy has been used and the corresponding capacity regions have been characterized. The disadvantages of weak secrecy are well known. The ultimate performance criteria for the key are perfect secrecy together with uniform distribution of the key. We derive the corresponding capacity region. We show that, for perfect secrecy and uniform distribution of the key, we can achieve the same rates as for weak secrecy together with a weaker requirement on the distribution of the key. In the classical works on PUF based authentication, it is assumed that the source statistics are known perfectly. This requirement is rarely met in applications. That is why the model is generalized to a compound model, taking into account source uncertainty. We also derive the capacity region for the compound model requiring perfect secrecy. Additionally, we consider results for secure storage using a biometric or PUF source that follow directly from the results for authentication. We also generalize known results for this problem by weakening the assumption concerning the distribution of the data that shall be stored. This allows us to combine source compression and secure storage.

Keywords: authentication; secure storage; perfect secrecy; privacy leakage

1. Introduction

The present work addresses two essential practical problems concerning secrecy in information systems. The first problem is authentication in order to manage access to the system. The second problem is secure storage in public databases. Both problems are of essential importance for further development of future communication systems. The goal of this work is to derive a fundamental characterization of the possible performance of such communication systems that meets very strict secrecy requirements. We show that these strict requirements can be met without loss in performance compared to known results with weaker secrecy requirements.

Information theoretic security has become a very active field of research in information theory in the past ten years, with a large number of promising approaches. For a current presentation, see [1]. In [2], the paper first introducing information theoretic security, the authors suggest requiring perfect secrecy [3] to guarantee security in communication. This means the data available to an attacker should be stochastically independent of the message that should be kept secret (the data and the message are modeled using random variables (RVs)). Thus, an attacker does not benefit from learning these data. In [4], this notion of security is weakened. The authors use weak secrecy [3] instead of perfect secrecy to guarantee secure communication. In many of the works on information theoretic security following [4],

one considers weak secrecy or strong secrecy [3], which is yet another security requirement that is also weaker than perfect secrecy. As the name suggests, perfect secrecy is the desired ideal situation in cryptographic applications where an attacker does not get any information about the secret. Considering the roots of information theoretic security and its intuitive motivation, it suggests itself to require perfect secrecy for secure communications. Additionally, in [3], the recommendation is to not use weak secrecy as a secrecy measure. In [5], there is an example of a protocol that is obviously not secure, but meets the weak secrecy requirement.

The authors of the landmark paper [6] derive the capacity for secret key generation requiring perfect secrecy. A different model in information theoretic security has as an essential feature a biometric source or a PUF source. The outputs of biometric sources and the outputs of PUF sources both uniquely characterize a person [7], or a device, respectively [8]. This property qualifies them for being used for authentication as well as for secure storage. In [7,9], the authors consider a model for authentication using the output of a biometric source. They also consider a model that can be interpreted as a model for secure storage using a biometric source. Both of these models are very similar to the model for secret key generation and for both of the models the authors require weak secrecy to hold when defining achievability.

In [6,7,9], the authors assume that the statistics of the (PUF) source are perfectly known. A simple analysis of [6,7,9] shows that the protocols for authentication constructed there heavily depend on the knowledge of the source statistics. Particularly, it is possible that small variations of the source statistics influence the reliability and secrecy of the protocols for authentication or storage, respectively. The assumption that the source statistics are perfectly known is too optimistic in applications. That is why we are interested in considering the uncertainty of the source or PUF source. We assume that we do not know the statistics of the source, but that we know a set of source statistics that contains the actual source statistic. Thus, we consider a compound version of the source model. We want to develop robust protocols that work for all source statistics in a given set. The compound model also allows us to describe an attack scenario where the attacker is able to alter the source statistics. There are relatively few results concerning compound sources. The compound version of the source model from [6] is considered in [10].

One of our contributions in the present work is the generalization of the model for authentication from [7], by considering authentication using a compound PUF source (or equivalently a biometric source). Additionally, our work differs from the state of the art as we consider protocols for authentication that achieve perfect secrecy.

We also consider secure data storage making use of a PUF source (or equivalently a biometric source). The corresponding information theoretic model is very similar to the second model presented in [7], but, in contrast to [7], we define achievability requiring perfect secrecy and we consider source uncertainty of the PUF source. Our considerations concerning perfect secrecy in this work answer the question posed in the conclusion of [11].

Some of the results for secure authentication described in this work have already been published in [12]. Here, we additionally present the proofs that have been omitted in [12], i.e., the proofs of Theorem 4 and Theorem 5 and some more discussion. The results concerning secure storage have been presented in [13,14]. As these results heavily depend on [12], we briefly state them here (as well as the corresponding definitions).

In Section 2, we describe the authentication process and define the corresponding information theoretic model. We discuss different definitions of achievability for the model in Section 3. In this context, protocols that achieve perfect secrecy are of special interest. We develop the corresponding definition of achievability in this section. In Section 4, we prove capacity results for the model with respect to the various definitions of achievability. The main result in this section is Theorem 2. In Section 5, we generalize the model for authentication to the case with source uncertainty and define achievability for this model in Section 6. In Section 7, we derive the capacity region for the compound storage model. In Section 8, we consider some results for secure storage that follow from our results for authentication.

The key result from authentication that we use for secure storage with perfect secrecy is Theorem 2. In Section 9, we further discuss our results.

For the most part, we use the notation introduced in [3].

2. Authentication Model

At first, we consider authentication using biometric or PUF data. This means we consider a scenario where a user enrolls in a system by giving a certain amount of biometric or PUF data to the system. Later, when the user wants to be authenticated, he again gives biometric or PUF data to the system. The system then decides if the user is accepted, i.e., if it is the same user that is enrolled in the system. In our considerations, we assume that the system can store some data in a public database.

Figure 1 depicts the authentication process as described in [7]. The process consists of two phases. In the first phase, the enrollment phase, the authentication system receives X^n from the PUF source and the ID of a user. It generates a helper message M and a secret-key K from X^n. It then uses a one-way function f on K and stores the result and M in a public database together with the user's ID. The second phase is the authentication phase. In this phase, the system receives Y^n from the PUF source and the ID of a user. It reads the corresponding helper message M and $f(K)$ from the database. From M and Y^n, it generates a secret-key $\hat{K}$. Then, the system compares $f(K)$ and $f(\hat{K})$. If they are equal, the user is accepted; otherwise, the user is rejected.

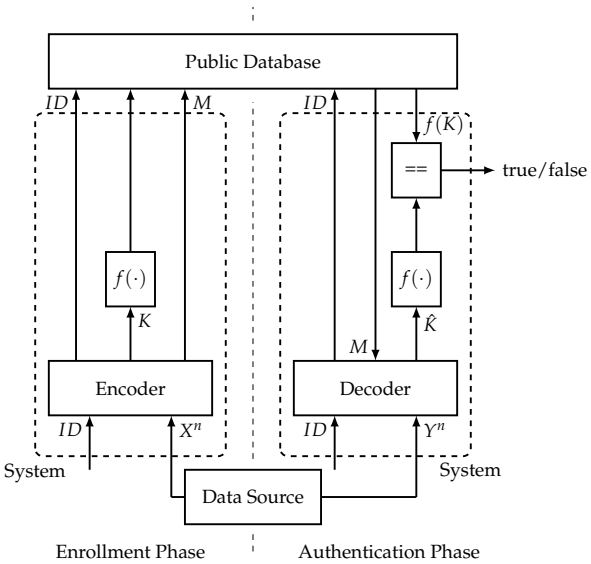

Figure 1. Authentication process considered in [7].

Now, we define an information theoretic model of the authentication process. We use random variables (RVs) to model the data. In the first chapters of this work, we assume that the distribution of the RVs is perfectly known. We drop this assumption in Section 5.

Definition 1. *Let $n \in \mathbb{N}$. The* authentication model *consists of a discrete memoryless multiple source (DMMS) with generic variables XY [3], the (possibly randomized) encoders [3] $\Phi\colon \mathcal{X}^n \to \mathcal{M}$, $\Theta\colon \mathcal{X}^n \to \mathcal{K}$ and the deterministic decoder $\psi\colon \mathcal{Y}^n \times \mathcal{M} \to \hat{\mathcal{K}}$. Let X^n and Y^n be the output of the DMMS. The RVs M and K are generated from X^n using Φ and Θ. The RV $\hat{K}$ is generated from Y^n and M using ψ. We use the term authentication protocol for (Φ, Θ, ψ).*

Remark 1. *It is possible to define the authentication protocol in a more general way by permitting randomized decoders Ψ, but one can argue that in our definition of achievability a randomized Ψ does not improve the performance of the protocols ([3], Problem 17.11). For convenience, we use the less general definition.*

Remark 2. *We model the PUF source as a DMMS. Due to physically induced distortions, we model the biometric/PUF data read in the two phases as jointly distributed RVs.*

Remark 3. *The distribution of XY is assumed to be known and can be used for the generation of the RVs. Thus, the encoders and the decoder are allowed to depend on the distribution.*

3. Various Definitions of Achievability

For the authentication model, we define achievable secret-key rate versus privacy-leakage rate pairs. Intuitively, we want the probability that a legitimate user is rejected in the authentication phase to be small. Thus, $\Pr(K = \hat{K})$ should be large to fulfill this reliability condition. Additionally, the probability that an attacker is accepted in the authentication phase should be as small as possible. Thus, we consider the maximum false acceptance probability (mFAP) [15], which is the probability that an attacker using the best possible attack strategy is accepted in the authentication phase averaged over all public messages $m \in \mathcal{M}$. As we want the mFAP to be as small as possible, we are interested in the largest possible set of secret keys $\mathcal{K}$. This reasoning is explained below. The system uses the output of a PUF source as input so it should leak as little information about X^n as possible [7]. This motivates the following definition of achievable rate pairs.

Definition 2. *A tuple (R, L), $R, L \geq 0$, is an* achievable *secret-key rate versus privacy-leakage rate pair for the authentication model if for every $\delta > 0$ there is an $n_0 = n_0(\delta)$ such that for all $n \geq n_0$ there exists an authentication protocol such that*

$$\begin{aligned} Pr(K = \hat{K}) &\geq 1 - \delta, \\ mFAP &\leq \tfrac{1}{|\mathcal{K}|}, \\ \tfrac{1}{n} \log |\mathcal{K}| &\geq R - \delta, \\ \tfrac{1}{n} I(M; X^n) &\leq L + \delta. \end{aligned} \tag{1}$$

We denote the corresponding authentication protocols by FAP-Protocols (False-Acceptance-Probability-Protocols).

Remark 4. *In [15], a very similar definition of achievability is used. Instead of considering the relation between the mFAP and the set of secret-keys* (1)*, the authors define the false-acceptance exponent that describes the exponential decrease of the mFAP in n. A rate pair (R, L) that is achievable using FAP-protocols is also achievable according to the definition in [15], R playing the role of the false-acceptance exponent.*

We now clarify the bound on the mFAP in Inequality (1) and our interest in large secret-key rates. For this purpose, we consider the following observation.

Lemma 1. *For a communication protocol fulfilling the reliability condition, it holds that*

$$mFAP \geq \tfrac{1-\delta}{|\mathcal{K}|}.$$

Proof. Introduce the RV E, setting $E = 1$ for $K \neq \hat{K}$ and $E = 0$, otherwise. Thus,

$$\begin{aligned}
\text{mFAP} &= \sum_{m\in\mathcal{M}} P_M(m) \max_{y^n\in\mathcal{Y}^n} P_{K|M}(\psi(y^n,m)|m) \\
&\geq \sum_{m\in\mathcal{M}} P_M(m) \max_{y^n\in\mathcal{Y}^n} P_{K|ME}(\psi(y^n,m)|m,0) P_{E|M}(0|m) \\
&\overset{(a)}{=} \sum_{m\in\mathcal{M}} P_{ME}(m,0) \max_{k\in\mathcal{K}} P_{K|ME}(k|m,0) \\
&\geq \sum_{m\in\mathcal{M}} P_{ME}(m,0) \tfrac{1}{|\mathcal{K}|} \\
&\overset{(b)}{\geq} (1-\delta)\tfrac{1}{|\mathcal{K}|}.
\end{aligned}$$

Here, (a) follows as $P_{K|ME}(k|m,0) = 0$ if there is no $y^n \in \mathcal{Y}^n$ such that $\psi(y^n,m) = k$ and (b) follows from the δ-recoverability of K from $\hat{K}$. □

Thus, Lemma 1 shows that requiring Inequality (1) is in fact equivalent to requiring the mFAP to be as small as possible. It also justifies our interest in a large set $\mathcal{K}$.

There is another way to define achievable secret-key rate versus privacy-leakage rate pairs for the authentication model. Here, we want to keep the key secret from the attacker. $H(K|M)$ can be interpreted as the average information required to specify k when m is known ([16], Chapter 2). Thus, we want $H(K|M)$ to be as large as possible instead of requiring a small mFAP. This means we require $\log|\mathcal{K}| = H(K|M)$. This condition is equivalent to the combination of the perfect secrecy condition $I(K;M) = 0$ [5] and the uniform distribution of the key, i.e., $H(K) = \log|\mathcal{K}|$. Thus, we define achievability as follows.

Definition 3. *A tuple (R,L), $R,L \geq 0$, is an* achievable *secret-key rate versus privacy-leakage rate pair for the authentication model if for every $\delta > 0$ there is an $n_0 = n_0(\delta)$ such that for all $n \geq n_0$ there exists an authentication protocol such that*

$$\begin{aligned}
Pr(K = \hat{K}) &\geq 1-\delta, \\
H(K) &= \log|\mathcal{K}|, &(2)\\
I(M;K) &= 0, &(3)\\
\tfrac{1}{n}\log|\mathcal{K}| &\geq R-\delta, \\
\tfrac{1}{n} I(M;X^n) &\leq L+\delta.
\end{aligned}$$

We denote the corresponding authentication protocols by PSA-Protocols *(Perfect-Secrecy-Authentication-Protocols).*

Remark 5. *In [6], the authors derive the secret-key capacity for the source model. They define achievability requiring perfect secrecy and uniform distribution of the key. They do not consider the privacy-leakage in contrast to our definition of achievability.*

It is interesting to compare the rate pairs achievable with respect to the restrictive Definition 3 with commonly used weaker requirements. In ([7], Definition 3.1), the authors give a different definition of achievable secret-key rate versus privacy-leakage rate pairs. Instead of Eqation (2), they require

$$H(K) \geq \log|\mathcal{K}| - \delta$$

and instead of Equation (3) they require

$$\tfrac{1}{n} I(M;K) \leq \delta,$$

which is called the weak secrecy condition [5]. Thus, we get a third definition of achievability.

Definition 4 ([7]). *A tuple (R, L), $R, L \geq 0$, is an* achievable *secret-key rate versus privacy-leakage rate pair for the authentication model if for every $\delta > 0$ there is an $n_0 = n_0(\delta)$ such that for all $n \geq n_0$ there exists an authentication protocol such that*

$$\begin{aligned} Pr(K = \hat{K}) &\geq 1 - \delta, \\ H(K) &\geq \log |\mathcal{K}| - \delta, \\ \tfrac{1}{n} I(M;K) &\leq \delta, \\ \tfrac{1}{n} \log |\mathcal{K}| &\geq R - \delta, \\ \tfrac{1}{n} I(M;X^n) &\leq L + \delta. \end{aligned}$$

We denote the corresponding authentication protocols by WSA-Protocols *(Weak-Secrecy-Authentication-Protocols).*

Definition 5. *The set of achievable rate pairs that are achievable using PSA-Protocols is called the capacity region $\mathcal{R}_{PSA}$. The set of achievable rate pairs that are achievable using WSA-Protocols is called the capacity region $\mathcal{R}_{WSA}$ and the set of achievable rate pairs that are achievable using FAP-Protocols is called the capacity region $\mathcal{R}_{FAP}$.*

Now, we look at some straightforward relations between these capacity regions. We can directly see that Definition 3 is more restrictive than Definition 4 so a PSA-Protocol is also a WSA-Protocol and thus

$$\mathcal{R}_{PSA} \subset \mathcal{R}_{WSA}. \tag{4}$$

We now show that a PSA-Protocol is also a FAP-Protocol.

Lemma 2. *It holds that*

$$\mathcal{R}_{PSA} \subset \mathcal{R}_{FAP}.$$

Proof. As Equations (2) and (3) imply, $P_{K|M}(k|m) = \frac{1}{|\mathcal{K}|}$ for all $(k, m) \in \mathcal{K} \times \mathcal{M}$, we have

$$\begin{aligned} \text{mFAP} &= \sum_{m \in \mathcal{M}} P_M(m) \max_{y^n \in \mathcal{Y}^n} P_{K|M}(\psi(y^n, m)|m) \\ &\leq \sum_{m \in \mathcal{M}} P_M(m) \max_{k \in \mathcal{K}} P_{K|M}(k|m) = \tfrac{1}{|\mathcal{K}|}. \end{aligned}$$

□

4. Capacity Regions for the Authentication Model

In ([7], Theorem 3.1), the authors derive the capacity region $\mathcal{R}_{WSA}$.

Theorem 1 ([7]). *It holds that*

$$\mathcal{R}_{WSA} = \bigcup_{U} \{(R, L) \colon 0 \leq R \leq I(U;Y), L \geq I(U;X) - I(U;Y)\}.$$

The union is over all RVs U such that $U - X - Y$. We only have to consider RVs U with $|\mathcal{U}| \leq |\mathcal{X}| + 1$.

Remark 6. *The authors of [7] do not consider randomized encoders. In contrast, we permit randomization of the encoders in the enrollment phase. Using the strategy described in ([3], Problem 17.15), one can use the converse for deterministic encoders to prove the converse for randomized encoders with the same bounds on the secret-key rate and the privacy-leakage rate. Thus, the converse in [7] also holds true when randomization is permitted.*

The following theorem is one of our main results.

Theorem 2. *It holds that*

$$\mathcal{R}_{PSA} = \mathcal{R}_{WSA}.$$

Proof. We do not prove Theorem 2 here but prove a more general result in the remainder of the text. This result is Theorem 5. It is more general as it is concerned with a compound version of the authentication model. The authentication model is a special case of the compound authentication model where the compound set consists of a single DMMS. □

We now strengthen Lemma 2.

Theorem 3. *It holds that*

$$\mathcal{R}_{PSA} = \mathcal{R}_{FAP}.$$

Proof. The achievability result is implied by Lemma 2. For the converse, we use a result of [15]. As discussed in Remark 4, a rate pair (R, L), which is achievable according to Definition 2 is also achievable according to the definition of achievability used in [15], where R plays the role of the false acceptance exponent E. Thus, we use ([15], Theorem 4), which says that a rate pair $(E, L) \notin \mathcal{R}_{WSA}$ is not achievable. This implies our converse. □

5. Compound Authentication Model

We now consider authentication when the data source is not perfectly known. Figure 2 shows the corresponding authentication process. The only difference to the authentication process in Section 2 is the source uncertainty. As one can see in Figure 2, we even assume that an attacker can influence the source in the sense that the state of the source is altered, i.e., it generates another statistic. If the protocol for authentication is not robust, then authentication will not work.

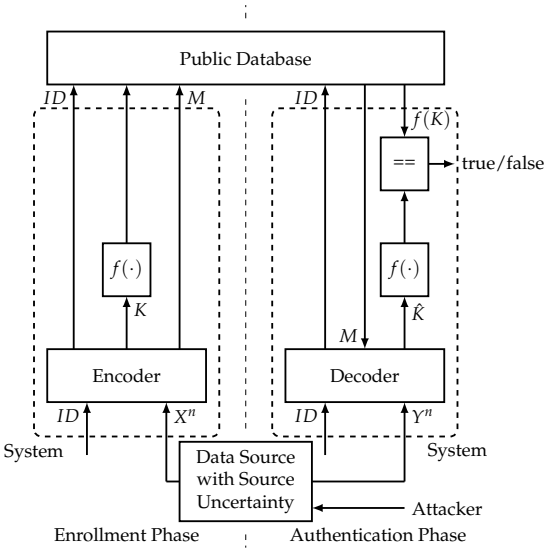

Figure 2. Authentication process with source uncertainty (as considered in [12]).

We define the following information theoretic model for this authentication process with source uncertainty.

Definition 6. *Let $n \in \mathbb{N}$. The compound authentication model consists of a set $\mathfrak{S}$ of DMMSs with generic variables X_sY_s, $s \in \mathcal{S}$, (all on the same alphabets $\mathcal{X}$ and $\mathcal{Y}$), the (possibly randomized) encoders $\Phi\colon \mathcal{X}^n \to \mathcal{M}$, $\Theta\colon \mathcal{X}^n \to \mathcal{K}$ and the (possibly randomized) decoder $\Psi\colon \mathcal{Y}^n \times \mathcal{M} \to \hat{\mathcal{K}}$. Let X^n and Y^n be the output of one of the DMMSs in $\mathfrak{S}$, i.e., $P_{XY} = P_{X_sY_s}$ for an $s \in \mathcal{S}$, but s is not known. The RVs M and K are generated from X^n using Φ and Θ. The RV $\hat{K}$ is generated from Y^n and M using Ψ. We use the term compound authentication protocol for (Φ, Θ, Ψ).*

Remark 7. *The uncertainty of the data source is modeled making use of a compound DMMS, that is, the DMMS modeling the PUF source is not known, but we know a set of DMMSs to which the actual DMMS belongs.*

Remark 8. *$\mathfrak{S}$ is assumed to be known and can be used for the generation of the RVs, that is, the encoder and the decoder can depend on these distributions.*

Definition 7. *Given $\mathfrak{S}$, we define the set*

$$\mathcal{I}(\hat{s}) = \{s \in \mathcal{S}\colon \sum_{y \in \mathcal{Y}} P_{X_sY_s}(x,y) = P_{X_{\hat{s}}}(x) \quad \forall x \in \mathcal{X}\}$$

for $\hat{s} \in \mathcal{S}$. The sets $\mathcal{I}(\hat{s})$, $\hat{s} \in \mathcal{S}$, form a partition of $\mathcal{S}$, as they form the equivalence classes for the corresponding equivalence relation. We denote a set of representatives by $\hat{\mathcal{S}}$.

6. Achievability for the Compound Model

For the compound authentication model, we define achievable secret-key rate versus privacy-leakage rate pairs.

Definition 8. *A tuple (R, L), $R, L \geq 0$, is an* achievable *secret-key rate versus privacy-leakage rate pair for the compound authentication model if for every $\delta > 0$ there is an $n_0 = n_0(\delta)$ such that, for all $n \geq n_0$, there exists a compound authentication protocol such that, for all $s \in \mathcal{S}$,*

$$Pr(K = \hat{K}) \geq 1 - \delta, \tag{5}$$
$$H(K) = \log|\mathcal{K}|, \tag{6}$$
$$I(M;K) = 0, \tag{7}$$
$$\tfrac{1}{n}\log|\mathcal{K}| \geq R - \delta,$$
$$\tfrac{1}{n} I(M;X^n) \leq L + \delta,$$

where $P_{XY} = P_{X_sY_s}$. We denote the corresponding authentication protocols by PSCA-Protocols (Perfect-Secrecy-Compound-Authentication-Protocols).

Definition 9. *The set of achievable secret-key versus privacy-leakage rate pairs that are achievable using PSCA-Protocols is called the compound capacity region $\mathcal{R}_{PSCA}(\mathfrak{S})$.*

7. Capacity Regions for the Compound Authentication Model

We now derive the compound capacity region $\mathcal{R}_{PSCA}(\mathfrak{S})$ for the compound authentication model. We only consider compound sets $\mathfrak{S}$ such that $|\hat{\mathcal{S}}| < \infty$. For the proof, we need the following theorem, which is a generalization of ([3], Theorem 6.10).

Theorem 4. *Given a (possibly infinite) set $\mathcal{W}$ of channels $W\colon \mathcal{X} \to \mathcal{Y}$, a set $A \subset \mathcal{X}^n$ with $P^n(A) > \eta$, $P \in \mathcal{P}(\mathcal{X})$, $\eta > 0$ and $\epsilon > 0$. Then, for every $\tau > 0$ and all n large enough, there is a pair of mappings (f, ϕ), $f\colon \mathcal{M}_f \to \mathcal{X}^n$, $\phi\colon \mathcal{Y}^n \to \mathcal{M}_f$, such that (f, ϕ) is an (n, ϵ)-code for all $W \in \mathcal{W}$ with codewords in A and*

$$\tfrac{1}{n} \log |\mathcal{M}_f| \geq \inf_{W \in \mathcal{W}} I(P; W) - \tau.$$

We call this pair of mappings a compound (n, ϵ)-code for $\mathcal{W}$.

Even though the proof of Theorem 4 is very similar to the proof of ([3], Theorem 6.10), the proof of ([17], Theorem 4.3) and the proof of the results in [18], we prove Theorem 4 for the sake of completeness. The proof can be found in Appendix A.

Theorem 5. *It holds that*

$$\begin{aligned}\mathcal{R}_{PSCA}(\mathfrak{S}) &= \bigcap_{\hat{s} \in \hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \{(R, L)\colon 0 \leq R \leq \inf_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s), L \geq \sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_s) - I(U_{\hat{s}}; Y_s)\} \\ &\overset{(a)}{=} \bigcap_{\hat{s} \in \hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \mathcal{R}_{\hat{s}}^{(PSCA)}(\mathfrak{S}, U_{\hat{s}}),\end{aligned}$$

where, for (a), we define $\mathcal{R}_{\hat{s}}^{(PSCA)}(\mathfrak{S}, U_{\hat{s}})$ appropriately. For all $\hat{s} \in \hat{\mathcal{S}}$, the union is over all RVs $U_{\hat{s}}$ such that, for all $s \in \mathcal{I}(\hat{s})$, we have $U_{\hat{s}} - X_s - Y_s$. For $|\mathcal{S}| < \infty$, we only have to consider RVs $U_{\hat{s}}$ with $|\mathcal{U}_{\hat{s}}| \leq |\mathcal{X}| + |\mathcal{I}(\hat{s})|$.

Proof. For all $\hat{s} \in \hat{\mathcal{S}}$ and all $s \in \mathcal{I}(\hat{s})$, let $U_{\hat{s}}$, X_s and Y_s be RVs where $X_s Y_s$ are the output of the DMMS in $\mathfrak{S}$ with index s and X_s and $U_{\hat{s}}$ are connected by the channel $V_{\hat{s}}\colon \mathcal{X} \to \mathcal{U}_{\hat{s}}$. Thus, we have the Markov chains $U_{\hat{s}} - X_s - Y_s$ for all $s \in \mathcal{I}(\hat{s})$. Let $\mathcal{U} = \bigcup_{\hat{s} \in \hat{\mathcal{S}}} \mathcal{U}_{\hat{s}}$. We now show that, given $\delta > 0$, for n large enough we can choose a set $\mathcal{C} \subset \mathcal{U}^n$ that consists of $|\mathcal{M}|$ disjoint subsets $\mathcal{C}_m$ with the following properties.

- We consider a partition of the set of all sets $\mathcal{C}_m$ in $|\hat{\mathcal{S}}|$ subsets. Thus, we denote the sets $\mathcal{C}_m$ by $\mathcal{C}_{m,\hat{s}}$, $\hat{s} \in \hat{\mathcal{S}}$, indicating to which subset they belong. We denote the set of indices m corresponding to $\hat{s}$ by $\mathcal{M}_{\hat{s}}$. For each $\mathcal{C}_{m,\hat{s}}$, we have

$$|\mathcal{C}_{m,\hat{s}}| = \lceil \inf_{s \in \mathcal{I}(\hat{s})} \exp\left(n(I(U_{\hat{s}}; Y_s) - \delta)\right) \rceil.$$

- Each $\mathcal{C}_{m,\hat{s}}$ consists of sequences of the same type.
- It holds that

$$P_{U_{\hat{s}}}^n(\mathcal{C}) > 1 - \eta \tag{8}$$

  for $\eta > 0$ and all $\hat{s} \in \hat{\mathcal{S}}$.
- For each $\hat{s} \in \hat{\mathcal{S}}$, one can define pairs of mappings that are compound (n, ϵ)-codes, $\epsilon > 0$, for the channels $W_s\colon \mathcal{U} \to \mathcal{Y}$, $W_s = P_{Y_s|U_{\hat{s}}}$ for all $s \in \mathcal{I}(\hat{s})$ in the following way. Define an (arbitrary) bijective mapping $f_m\colon \{1 \cdots |\mathcal{C}_{m,\hat{s}}|\} \to \mathcal{C}_{m,\hat{s}}$ and an appropriate mapping $\phi_m\colon \mathcal{Y}^n \to \{1 \cdots |\mathcal{C}_{m,\hat{s}}|\}$. Then, (f_m, ϕ_m) is such a code. This means

$$W_s^n(\phi_m^{-1}(f_m^{-1}(u^n))|u^n) \geq 1 - \epsilon \tag{9}$$

  for all $s \in \mathcal{I}(\hat{s})$ and for all codewords u^n in $\mathcal{C}_{m,\hat{s}}$. This is possible for all $m \in \mathcal{M}_{\hat{s}}$.

Let $\delta' > 0$. We denote the elements of $\hat{\mathcal{S}}$ by $\hat{s}_1, \hat{s}_2, \cdots, \hat{s}_{|\hat{\mathcal{S}}|}$. We consider $\mathcal{T}^n_{P_{U_{\hat{s}_1}}, \xi}, \mathcal{T}^n_{P_{U_{\hat{s}_2}}, \xi}, \cdots, \mathcal{T}^n_{P_{U_{\hat{s}_{|\hat{\mathcal{S}}|}}}, \xi}$, $\xi > 0$, which are disjoint subsets of $\mathcal{U}^n$. We show that they are in fact disjoint subsets of $\mathcal{U}^n$ for ξ small

enough. This can be seen as follows. For $\hat{s}_i, \hat{s}_j \in \hat{\mathcal{S}}$, $\hat{s}_i \neq \hat{s}_j$, it holds that $P_{U_{\hat{s}_i}}(u) \neq P_{U_{\hat{s}_j}}(u)$ for at least one $u \in \mathcal{U}$. Thus, there is a $u \in \mathcal{U}$ with

$$|P_{U_{\hat{s}_i}}(u) - P_{U_{\hat{s}_j}}(u)| > \alpha$$

for some $\alpha > 0$.

Now, assume that there is a $u^n \in \mathcal{T}^n_{P_{U_{\hat{s}_i}},\xi} \cap \mathcal{T}^n_{P_{U_{\hat{s}_j}},\xi}$. Denote the type of u^n by p_{u^n}. Thus, there is a $u \in \mathcal{U}$ with

$$\begin{aligned}\alpha &< |P_{U_{\hat{s}_i}}(u) - P_{U_{\hat{s}_j}}(u)| \\ &= |P_{U_{\hat{s}_i}}(u) - P_{u^n}(u) + P_{u^n}(u) - P_{U_{\hat{s}_j}}(u)| \\ &\leq |P_{U_{\hat{s}_i}}(u) - P_{u^n}(u)| + |P_{U_{\hat{s}_j}}(u) - P_{u^n}(u)| \leq 2\xi,\end{aligned}$$

where the last inequality follows from the assumption that $u^n \in \mathcal{T}^n_{P_{U_{\hat{s}_i}},\xi} \cap \mathcal{T}^n_{P_{U_{\hat{s}_j}},\xi}$. Thus, for $\xi < \alpha/2$, this is a contradiction and we know $\mathcal{T}^n_{P_{U_{\hat{s}_i}},\xi}$ and $\mathcal{T}^n_{P_{U_{\hat{s}_j}},\xi}$ are disjoint.

We start the construction of $\mathcal{C}$ by choosing a set $\mathcal{A}_{1,\hat{s}_1} \subset \mathcal{T}^n_{P_{U_{\hat{s}_1}},\xi}$ with $P^n_{U_{\hat{s}_1}}(\mathcal{A}_{1,\hat{s}_1}) \geq \eta'$ with $\eta > \eta' > 0$. According to Theorem 4, there is a compound (n,ϵ)-code for the channels W_s, $s \in \mathcal{I}(\hat{s}_1)$ with at least

$$\lceil \inf_{s\in\mathcal{I}(\hat{s}_1)} \exp\left(n(I(U_{\hat{s}_1};Y_s) - \delta')\right) \rceil$$

codewords $u^n \in \mathcal{A}_{1,\hat{s}_1}$ for n large enough. We denote the set of these codewords by $\mathcal{C}'_{1,\hat{s}_1}$. As there are less than $(n+1)^{|\mathcal{U}|}$ types, we know that there is a set of at least

$$\left\lceil \frac{\lceil \inf_{s\in\mathcal{I}(\hat{s}_1)} \exp\left(n(I(U_{\hat{s}_1};Y_s)-\delta')\right)\rceil}{(n+1)^{|\mathcal{U}|}} \right\rceil$$

codewords in $\mathcal{C}'_{1,\hat{s}_1}$ with the same type. We only pick these codewords. There are at least

$$\left\lceil \inf_{s\in\mathcal{I}(\hat{s}_1)} \exp\left(n(I(U_{\hat{s}_1};Y_s) - \delta' - |\mathcal{U}|\tfrac{\log(n+1)}{n})\right)\right\rceil \geq \lceil \inf_{s\in\mathcal{I}(\hat{s}_1)} \exp\left(n(I(U_{\hat{s}_1};Y_s) - \delta)\right)\rceil$$

of them for n large enough. We now pick exactly

$$\lceil \inf_{s\in\mathcal{I}(\hat{s}_1)} \exp\left(n(I(U_{\hat{s}_1};Y_s) - \delta)\right)\rceil$$

of these codewords and we denote this set by $\mathcal{C}_{1,\hat{s}_1}$. Now, we choose a set $\mathcal{A}_{2,\hat{s}_1} \subset \mathcal{T}^n_{P_{U_{\hat{s}_1}},\xi} \setminus \mathcal{C}_{1,\hat{s}_1}$ with $P^n_U(\mathcal{A}_{2,\hat{s}_1}) \geq \eta'$. We construct the set $\mathcal{C}_{2,\hat{s}_1}$ in the same way as $\mathcal{C}_{1,\hat{s}_1}$. Thus, $\mathcal{C}_{2,\hat{s}_1}$ is a set of

$$\lceil \inf_{s\in\mathcal{I}(\hat{s}_1)} \exp\left(n(I(U_{\hat{s}_1};Y_s) - \delta)\right)\rceil$$

codewords of the same type corresponding to an (n,ϵ)-code. We continue this process until we can not find a set

$$\mathcal{A}_{|\mathcal{M}_{\hat{s}_1}|+1,\hat{s}_1} \subset \mathcal{T}^n_{P_{U_{\hat{s}_1}},\xi} \setminus \bigcup_{i\in\mathcal{M}_{\hat{s}_1}} \mathcal{C}_{i,\hat{s}_1}$$

with

$$P^n_{U_{\hat{s}_1}}(\mathcal{A}_{|\mathcal{M}_{\hat{s}_1}|+1,\hat{s}_1}) \geq \eta'.$$

This means

$$P^n_{U_{\hat{s}_1}}((\bigcup_{i\in\mathcal{M}_{\hat{s}_1}}\mathcal{C}_{i,\hat{s}_1})^c\cap\mathcal{T}^n_{P_{U_{\hat{s}_1}},\xi})<\eta'.$$

We repeat this process for all $\hat{s}\neq\hat{s}_1$, $\hat{s}\in\hat{\mathcal{S}}$. Thus, we have for all $\hat{s}\in\hat{\mathcal{S}}$

$$\begin{aligned}P^n_{U_{\hat{s}}}(\mathcal{C})&\geq P^n_{U_{\hat{s}}}(\bigcup_{i\in\mathcal{M}_{\hat{s}}}\mathcal{C}_{i,\hat{s}})\\&=1-P^n_{U_{\hat{s}}}((\bigcup_{i\in\mathcal{M}_{\hat{s}}}\mathcal{C}_{i,\hat{s}})^c)\\&=1-P^n_{U_{\hat{s}}}((\bigcup_{i\in\mathcal{M}_{\hat{s}}}\mathcal{C}_{i,\hat{s}})^c\cap\mathcal{T}^n_{P_{U_{\hat{s}}},\xi})-P^n_{U_{\hat{s}}}((\bigcup_{i\in\mathcal{M}_{\hat{s}}}\mathcal{C}_{i,\hat{s}})^c\cap(\mathcal{T}^n_{P_{U_{\hat{s}}},\xi})^c)\\&\geq 1-P^n_{U_{\hat{s}}}((\bigcup_{i\in\mathcal{M}_{\hat{s}}}\mathcal{C}_{i,\hat{s}})^c\cap\mathcal{T}^n_{P_{U_{\hat{s}}},\xi})-P^n_{U_{\hat{s}}}((\mathcal{T}^n_{P_{U_{\hat{s}}},\xi})^c).\end{aligned}$$

Thus, we have Inequality (8) for n large enough.
We now can define the encoders/decoders Φ, Θ and Ψ.

- We define Φ and Θ as follows. The system gets a sequence x^n. It checks if $x^n\in\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}$, $\xi'>0$, for an $\hat{s}\in\hat{\mathcal{S}}$ (We can choose ξ' small enough and n large enough such that the $\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}$ are disjoint). If this is true for $\hat{s}$, the channel $V_{\hat{s}}$ is used n times to generate u^n from x^n. For Φ, the system looks in $\mathcal{C}$ for u^n. If $u^n\in\mathcal{C}$ the system chooses for m the index of the subset $\mathcal{C}_m$ containing u^n. If $u^n\notin\mathcal{C}$ it chooses an arbitrary $m\in\mathcal{M}$. In addition, if $x^n\notin\bigcup_{\hat{s}\in\hat{\mathcal{S}}}\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}$, it chooses an arbitrary $m\in\mathcal{M}$. For Θ, the system looks in $\mathcal{C}$ for u^n. If $u^n\in\mathcal{C}$, it considers the compound (n,ϵ)-code corresponding to the subset $\mathcal{C}_{m,\hat{s}}$ containing u^n. If

$$|\mathcal{C}_{m,\hat{s}}|>\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil\inf_{s\in\mathcal{I}(\hat{s})}\exp\left(n(I(U_{\hat{s}};Y_s)-\delta)\right)\rceil,$$

we consider the following deterministic mapping $h_m\colon f_m^{-1}(\mathcal{C}_m)\to\mathcal{K}\cup\{\tilde{k}\}$. Here,

$$\mathcal{K}=\{1\cdots\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil\inf_{s\in\mathcal{I}(\hat{s})}\exp\left(n(I(U_{\hat{s}};Y_s)-\delta)\right)\rceil\}.$$

The preimage of any $k\in\mathcal{K}$ under h_m is a subset of $f_m^{-1}(\mathcal{C}_m)$ of size

$$\left\lfloor\frac{|\mathcal{C}_{m,\hat{s}}|}{\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil\inf_{s\in\mathcal{I}(\hat{s})}\exp(n(I(U_{\hat{s}};Y_s)-\delta))\rceil}\right\rfloor.$$

The rest of the $k'\in f_m^{-1}(\mathcal{C}_m)$ is mapped on $\tilde{k}\notin\mathcal{K}$. If

$$|\mathcal{C}_{m,\hat{s}}|=\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil\inf_{s\in\mathcal{I}(\hat{s})}\exp\left(n(I(U_{\hat{s}};Y_s)-\delta)\right)\rceil,$$

the system chooses $k=f_m^{-1}(u^n)$. In this case, we also define $h_m\colon f_m^{-1}(\mathcal{C}_m)\to\mathcal{K}\cup\{\tilde{k}\}$ where h_m is injective. If $u^n\notin\mathcal{C}$, k is chosen at random according to a uniform distribution on the alphabet. The same holds if u^n is mapped on $\tilde{k}$ or if $x^n\notin\bigcup_{\hat{s}\in\hat{\mathcal{S}}}\mathcal{T}^n_{P_{X,\hat{s}},\xi'}$.
- We define Ψ as follows. The system gets a sequence y^n and m. It decodes y^n using the code corresponding to $\mathcal{C}_{m,\hat{s}}$. Then, h_m is used on the result. The result is $\hat{k}$ if it differs from $\tilde{k}$. Otherwise, an arbitrary $\hat{k}\in\mathcal{K}$ is chosen.

Using the properties of the communication protocol, we analyse the achievability conditions. We denote the outputs of the DMMS by X^n and Y^n and the output of the channel used on X^n by U^n. Assume the index of the DMMS is $s\in\mathcal{I}(\hat{s})$, $\hat{s}\in\mathcal{S}$. Thus, $P_{X^nY^n}=P^n_{X_sY_s}$.

- We define the following events:

$$\begin{aligned}
\mathcal{E}_1 =& \{(x^n, y^n, u^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{U}^n \colon (x^n, u^n) \notin \mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''}\},\\
\mathcal{E}_2 =& \{(x^n, y^n, u^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{U}^n \colon u^n \notin \mathcal{C}\},\\
\mathcal{E}_3 =& \bigcup_{m \in \mathcal{M}} \{(x^n, y^n, u^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{U}^n \colon u^n \in \mathcal{C}_m \wedge h_m(f_m^{-1}(u^n)) = \tilde{k}\},\\
\mathcal{E}_4 =& \bigcup_{m \in \mathcal{M}} \{(x^n, y^n, u^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{U}^n \colon u^n \in \mathcal{C}_m \wedge f_m^{-1}(u^n) \neq \phi_m(y^n)\}.
\end{aligned}$$

According to ([3], Lemma 2.10), we can choose ξ'' small enough such that $(x^n, u^n) \in \mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''}$ implies $x^n \in \mathcal{T}^n_{P_{X_s}, \xi'}$ and $u^n \in \mathcal{T}^n_{P_{U_{\hat{s}}}, \xi}$. We have

$$\begin{aligned}
P_{X^n Y^n U^n}(\mathcal{E}_1) &= 1 - P_{X^n Y^n U^n}(\mathcal{E}_1^c)\\
&\overset{(a)}{=} 1 - P^n_{X_s Y_s U_{\hat{s}}}(\mathcal{E}_1^c) = 1 - P^n_{X_s U_{\hat{s}}}(\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})\\
&= P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c).
\end{aligned}$$

Here, (a) follows as for $x^n \in \mathcal{T}^n_{P_{X_s}, \xi'}$ the system uses $V_{\hat{s}}$ to generate u^n from x^n. Thus,

$$\begin{aligned}
\Pr(K \neq \hat{K}) &\leq P_{X^n Y^n U^n}(\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4)\\
&= P_{X^n Y^n U^n}(\mathcal{E}_1) + P_{X^n Y^n U^n}((\mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4) \cap \mathcal{E}_1^c)\\
&= P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c) + P_{X^n Y^n U^n}(\mathcal{E}_2 \cap \mathcal{E}_1^c) + P_{X^n Y^n U^n}((\mathcal{E}_3 \cup \mathcal{E}_4) \cap \mathcal{E}_1^c \cap (\mathcal{E}_2^c \cup \mathcal{E}_1)).
\end{aligned}$$

Now, we use

$$\begin{aligned}
P_{X^n Y^n U^n}(\mathcal{E}_2 \cap \mathcal{E}_1^c) &\leq \sum_{\substack{(x^n, u^n)\colon\\ x^n \in \mathcal{T}^n_{P_{X_s}, \xi'}\\ \wedge u^n \in \mathcal{C}^c}} P_{X^n U^n}(x^n, u^n)\\
&= \sum_{\substack{(x^n, u^n)\colon\\ x^n \in \mathcal{T}^n_{P_{X_s}, \xi'}\\ \wedge u^n \in \mathcal{C}^c}} P^n_{X_s U_{\hat{s}}}(x^n, u^n)\\
&\leq \sum_{\substack{(x^n, u^n)\colon\\ x^n \in \mathcal{X}^n\\ \wedge u^n \in \mathcal{C}^c}} P^n_{X_s U_{\hat{s}}}(x^n, u^n) = P^n_{U_{\hat{s}}}(\mathcal{C}^c)
\end{aligned}$$

and get

$$\begin{aligned}
\Pr(K \neq \hat{K}) &\leq P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c) + P^n_{U_{\hat{s}}}(\mathcal{C}^c) + P_{X^n Y^n U^n}((\mathcal{E}_3 \cup \mathcal{E}_4) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c)\\
&\leq P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c) + P^n_{U_{\hat{s}}}(\mathcal{C}^c) + P_{X^n Y^n U^n}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c) + P_{X^n Y^n U^n}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c).
\end{aligned}$$

Now, we define the RV $E = e(X^n, U^n)$ with $e\colon \mathcal{X}^n \times \mathcal{U}^n \to \{0, 1\}$

$$e(x^n, u^n) = \begin{cases} 0, & \text{for } u^n \in \mathcal{C} \wedge x^n \in \bigcup_{\hat{s} \in \hat{\mathcal{S}}} \mathcal{T}^n_{P_{X_{\hat{s}}}, \xi'},\\ 1, & \text{otherwise.} \end{cases}$$

We have

$$\begin{aligned}\Pr(K \neq \hat{K}) \leq\ & P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c) + P^n_{U_{\hat{s}}}(\mathcal{C}^c) \\ & + \sum_{m \in \mathcal{M}} P_M(m) P_{X^n Y^n U^n|M}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m) \\ & + \sum_{m \in \mathcal{M}} P_{ME}(m,0) P_{X^n Y^n U^n|ME}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m, 0)\end{aligned}$$

as for all $m \in \mathcal{M}$

$$P_{X^n Y^n U^n|ME}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m, 1) = 0.$$

As $u^n \in \mathcal{C}$ and $u^n \in \mathcal{T}^n_{P_{U_{\hat{s}}}, \xi}$ imply $u^n \in \mathcal{C}_m$ for an $m \in \mathcal{M}_{\hat{s}}$, we know

$$P_{X^n Y^n U^n|M}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m) = 0$$

and

$$P_{X^n Y^n U^n|M}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m) = 0$$

for $m \notin \mathcal{M}_{\hat{s}}$. Thus, we have

$$\begin{aligned}\Pr(K \neq \hat{K}) \leq\ & P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c) + P^n_{U_{\hat{s}}}(\mathcal{C}^c) \\ & + \sum_{m \in \mathcal{M}_{\hat{s}}} P_M(m) P_{X^n Y^n U^n|M}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m) \\ & + \sum_{m \in \mathcal{M}_{\hat{s}}} P_{ME}(m,0) P_{X^n Y^n U^n|ME}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m, 0).\end{aligned}$$

We know for $m \in \mathcal{M}_{\hat{s}}$

$$\begin{aligned}& P_{X^n Y^n U^n|M}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c | m) \\ & \leq \sum_{\substack{(x^n, y^n, u^n): \\ f_m^{-1}(u^n) \neq \phi_m(y^n) \\ \wedge u^n \in \mathcal{C}_m \wedge x^n \in \mathcal{T}^n_{P_{X_s}, \xi'}}} P_{X^n Y^n U^n|M}(x^n, y^n, u^n | m) \\ & = \sum_{\substack{(x^n, y^n, u^n): \\ f_m^{-1}(u^n) \neq \phi_m(y^n) \\ \wedge u^n \in \mathcal{C}_m \wedge x^n \in \mathcal{T}^n_{P_{X_s}, \xi'}}} P_{X^n|U^n Y^n M}(x^n | u^n, y^n, m) P_{Y^n|U^n M}(y^n | u^n, m) P_{U^n|M}(u^n | m).\end{aligned}$$

Using $M - U^n - Y^n$, we have

$$
\begin{aligned}
&P_{X^nY^nU^n|M}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c|m) \\
&\leq \sum_{\substack{(x^n,y^n,u^n):\\ f_m^{-1}(u^n)\neq\phi_m(y^n)\\ \wedge u^n\in\mathcal{C}_m\wedge x^n\in\mathcal{T}^n_{P_{X_s},\xi'}}} P_{X^n|U^nY^nM}(x^n|u^n,y^n,m)P_{Y^n|U^n}(y^n|u^n)P_{U^n|M}(u^n|m) \\
&= \sum_{\substack{(x^n,y^n,u^n):\\ f_m^{-1}(u^n)\neq\phi_m(y^n)\\ \wedge u^n\in\mathcal{C}_m\wedge x^n\in\mathcal{T}^n_{P_{X_s},\xi'}}} P_{X^n|U^nY^nM}(x^n|u^n,y^n,m)W_s^n(y^n|u^n)P_{U^n|M}(u^n|m) \\
&\leq \sum_{\substack{(x^n,y^n,u^n):\\ f_m^{-1}(u^n)\neq\phi_m(y^n)\\ \wedge u^n\in\mathcal{C}_m}} P_{X^n|U^nY^nM}(x^n|u^n,y^n,m)W_s^n(y^n|u^n)P_{U^n|M}(u^n|m) \\
&= \sum_{\substack{(y^n,u^n):\\ f_m^{-1}(u^n)\neq\phi_m(y^n)\\ \wedge u^n\in\mathcal{C}_m}} W_s^n(y^n|u^n)P_{U^n|M}(u^n|m) \\
&= \sum_{u^n\in\mathcal{C}_m} W_s^n((\phi_m^{-1}(f_m^{-1}(u^n)))^c|u^n)P_{U^n|M}(u^n|m).
\end{aligned}
$$

Thus, using Inequality (9), we have

$$\sum_{m\in\mathcal{M}_{\hat{s}}} P_M(m)P_{X^nY^nU^n|M}(\mathcal{E}_4 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c|m) \leq \epsilon$$

for n large enough. Now, consider $u^n \in \mathcal{C}_m$, $m \in \mathcal{M}$. We get

$$
\begin{aligned}
P_{U^n|ME}(u^n|m,0) &= \sum_{x^n\in\mathcal{X}^n} P_{U^nX^n|ME}(u^n,x^n|m,0) \\
&= \sum_{\hat{s}\in\hat{\mathcal{S}}}\sum_{x^n\in\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}} P_{U^nX^n|ME}(u^n,x^n|m,0)
\end{aligned}
$$

as

$$P_{U^nX^n|ME}(u^n,x^n|m,0) = 0$$

for $x^n \notin \bigcup_{\hat{s}\in\hat{\mathcal{S}}} \mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}$. We realize that, for $u^n \in \mathcal{C}_m$ and $x^n \in \bigcup_{\hat{s}\in\hat{\mathcal{S}}} \mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}$,

$$
\begin{aligned}
P_{U^nX^n|ME}(u^n,x^n|m,0) &= \tfrac{P_{U^nX^nME}(u^n,x^n,m,0)}{P_{ME}(m,0)} \\
&= \tfrac{P_{U^nX^n}(u^n,x^n)}{P_{ME}(m,0)}P_{ME|U^nX^n}(m,0|u^n,x^n) = \tfrac{P_{U^nX^n}(u^n,x^n)}{P_{ME}(m,0)},
\end{aligned}
$$

where the last step follows as

$$P_{ME|U^nX^n}(m,0|u^n,x^n) = 1.$$

Thus, we get

$$\begin{aligned} P_{U^n|ME}(u^n|m,0) &= \sum_{\hat{s}\in\hat{\mathcal{S}}} \sum_{x^n\in\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}} \frac{P_{U^nX^n}(u^n,x^n)}{P_{ME}(m,0)} \\ &= \sum_{\hat{s}\in\hat{\mathcal{S}}} \sum_{x^n\in\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}} \frac{P^n_{X_s}(x^n)V^n_{\hat{s}}(u^n|x^n)}{P_{ME}(m,0)} \\ &= \sum_{\hat{s}\in\hat{\mathcal{S}}} \sum_{\substack{p\in\mathcal{P}(n,\mathcal{X}):\\ |p(x)-p_{X_{\hat{s}}}(x)|\le\xi' \\ \forall x\in\mathcal{X}}} \sum_{x^n\in\mathcal{T}^n_p} \frac{\prod_{i=1}^n P_{X_s}(x_i)V_{\hat{s}}(u_i|x_i)}{P_{ME}(m,0)}. \end{aligned}$$

The last term is constant for all u^n of the same type. Thus,

$$P_{U^n|ME}(u^n|m,0) = p_{\mathcal{C}_m}$$

is constant for $u^n \in \mathcal{C}_m$. As

$$P_{U^n|ME}(u^n|m,0) = 0$$

for $u^n \notin \mathcal{C}_m$, we have

$$P_{U^n|ME}(u^n|m,0) = \tfrac{1}{|\mathcal{C}_m|}$$

for $u^n \in \mathcal{C}_m$. Now, we get

$$\begin{aligned} P_{X^nY^nU^n|ME}(\mathcal{E}_3\cup\mathcal{E}_1^c\cup\mathcal{E}_2^c|m,0) &\le \sum_{\substack{(x^n,y^n,u^n):\\ \wedge u^n\in\mathcal{C}_m\wedge x^n\in\mathcal{T}^n_{P_{X_{\hat{s}}},\xi'}\\ \wedge h_m(f_m^{-1}(u^n))=\tilde{k}}} P_{X^nY^nU^n|ME}(x^n,y^n,u^n|m,0) \\ &\le \sum_{\substack{u^n\in\mathcal{C}_m\\ \wedge h_m(f_m^{-1}(u^n))=\tilde{k}}} P_{U^n|ME}(u^n|m,0) = |h_m^{-1}(\tilde{k})|p_{\mathcal{C}_m}. \end{aligned}$$

We have

$$|h_m^{-1}(\tilde{k})| = |\mathcal{C}_m| - \min_{\hat{s}\in\hat{\mathcal{S}}} \lceil \inf_{s\in\mathcal{I}(\hat{s})} \exp\left(n(I(U_{\hat{s}};Y_s)-\delta)\right)\rceil \left\lfloor \frac{|\mathcal{C}_m|}{\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil \inf_{s\in\mathcal{I}(\hat{s})}\exp(n(I(U_{\hat{s}};Y_s)-\delta))\rceil} \right\rfloor$$

and get

$$\begin{aligned} &P_{X^nY^nU^n|ME}(\mathcal{E}_3\cup\mathcal{E}_1^c\cup\mathcal{E}_2^c|m,0) \\ &\le \tfrac{1}{|\mathcal{C}_m|}(|\mathcal{C}_m| - \min_{\hat{s}\in\hat{\mathcal{S}}} \lceil \inf_{s\in\mathcal{I}(\hat{s})} \exp\left(n(I(U_{\hat{s}};Y_s)-\delta)\right)\rceil (\tfrac{|\mathcal{C}_m|}{\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil \inf_{s\in\mathcal{I}(\hat{s})}\exp(n(I(U_{\hat{s}};Y_s)-\delta))\rceil} - 1)) \\ &= \frac{\min_{\hat{s}\in\hat{\mathcal{S}}}\lceil \inf_{s\in\mathcal{I}(\hat{s})}\exp(n(I(U_{\hat{s}};Y_s)-\delta))\rceil}{|\mathcal{C}_m|} \le \frac{2}{\exp(n\tilde{\epsilon})} \end{aligned}$$

or

$$P_{X^nY^nU^n|ME}(\mathcal{E}_3\cup\mathcal{E}_1^c\cup\mathcal{E}_2^c|m,0) = 0$$

respectively, if, for the source state s, it holds that $s \in \mathcal{I}(\hat{s})$ for the $\hat{s}$ corresponding to the smallest $\mathcal{C}_{m,\hat{s}}$. Here,

$$\tilde{\epsilon} = \inf_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) - \min_{\hat{s}'\in\mathcal{S}} \inf_{s\in\mathcal{I}(\hat{s}')} I(U_{\hat{s}'};Y_s).$$

Thus, for n large enough,

$$P_e \leq P^n_{X_s U_{\hat{s}}}((\mathcal{T}^n_{P_{X_s U_{\hat{s}}}, \xi''})^c) + \eta + \epsilon + \frac{2}{\exp(n\bar{\epsilon})}$$

and Inequality (5) is fulfilled for small enough constants and n large enough.

- We define $\tilde{k}\colon \mathcal{U}^n \times \mathcal{M} \to \{0,1\}$

$$\tilde{k}(u^n, m) = \begin{cases} 1, & \text{for } u^n \in f_m(h_m^{-1}(\tilde{k})), \\ 0, & \text{otherwise,} \end{cases}$$

and the RV $\tilde{K} = \tilde{k}(U^n, M)$. We have

$$P_{K|ME\tilde{K}}(k|m,0,0) = P_{U^n|ME\tilde{K}}(f_m(h_m^{-1}(k))|m,0,0).$$

Now, consider $u^n \in \mathcal{C}_m$. It holds that

$$P_{U^n|ME\tilde{K}}(u^n|m,0,0) = \frac{P_{U^n|ME}(u^n|m,0)}{P_{\tilde{K}|ME}(0|m,0)} P_{\tilde{K}|MEU^n}(0|m,0,u^n).$$

We know

$$P_{\tilde{K}|MEU^n}(0|m,0,u^n) = 1$$

for $u^n \notin f_m(h_m^{-1}(\tilde{k}))$. Thus,

$$P_{K|ME\tilde{K}}(k|m,0,0) = \frac{P_{U^n|ME}(h_m^{-1}(k)|m,0)}{P_{\tilde{K}|ME}(0|m,0)} = \frac{p_{\mathcal{C}_m}|h_m^{-1}(k)|}{P_{\tilde{K}|ME}(0|m,0)}$$

for all $k \in \mathcal{K}$. This means

$$P_{K|ME\tilde{K}}(k|m,0,0) = \frac{1}{|\mathcal{K}|},$$

as $|h_m^{-1}(k)|$ is constant for all $k \in \mathcal{K}$. We also know

$$H(K|M=m, E=e, \tilde{K}=\tilde{k}) = \log|\mathcal{K}|$$

for $P_{ME\tilde{K}}(m,e,\tilde{k}) > 0$, $(e,\tilde{k}) \neq (0,0)$ as k is chosen according to a uniform distribution on $\mathcal{K}$ in this case. Thus,

$$\begin{aligned} \log|\mathcal{K}| &\geq H(K|M) \geq H(K|ME\tilde{K}) \\ &= \sum_{\substack{(m,e,\tilde{k}) \\ \in \mathcal{M}\times\{0,1\}\times\{0,1\}}} P_{ME\tilde{K}}(m,e,\tilde{k}) H(K|M=m, E=e, \tilde{K}=\tilde{k}) = \log|\mathcal{K}|. \end{aligned}$$

This means Equations (6) and (7) are fulfilled.

- For the secret-key rate, we have

$$\frac{1}{n}\log|\mathcal{K}| \geq \min_{\hat{s}\in\hat{\mathcal{S}}} \inf_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - \delta. \tag{10}$$

- Finally, we analyse the privacy-leakage rate. We have

$$\begin{aligned} I(X^n; M) &= H(M) - H(M|X^n) - H(M|U^n) + H(M|U^n) \\ &= I(U^n; M) - H(M|X^n), \end{aligned}$$

where we use $H(M|U^n) = 0$ for the second equality (see ([3], Problem 3.1)). Now, we use

$$\begin{aligned} P_{ME}(\mathcal{M}_{\hat{s}}, 0) &\geq P_{X^nY^nU^n}(\mathcal{E}_1^c \cup \mathcal{E}_2^c) = P^n_{X_sY_sU_{\hat{s}}}(\mathcal{E}_1^c \cup \mathcal{E}_2^c) \\ &\geq P^n_{X_sY_sU_{\hat{s}}}(\mathcal{E}_1^c) + P^n_{X_sY_sU_{\hat{s}}}(\mathcal{E}_2^c) - 1 \\ &= P^n_{X_sU_{\hat{s}}}(\mathcal{T}^n_{P_{X_sU_{\hat{s}}},\xi''}) + P^n_{U_{\hat{s}}}(\mathcal{C}) - 1 \\ &\geq 1 - \eta - P^n_{X_sU_{\hat{s}}}((\mathcal{T}^n_{P_{X_sU_{\hat{s}}},\xi''})^c) \geq 1 - \zeta \end{aligned}$$

for $\zeta > 0$ and n large enough. We also use $P_{U^n|ME}(u^n|m, 0) = \frac{1}{|\mathcal{C}_m|}$ for $u^n \in \mathcal{C}_m$ and get

$$\begin{aligned} H(U^n|M) &\geq H(U^n|ME) \\ &\geq \sum_{m \in \mathcal{M}_{\hat{s}}} P_{ME}(m, 0) H(U^n|M = m, E = 0) \\ &\geq (1 - \zeta)(\min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - \delta)n. \end{aligned}$$

Thus,

$$I(X^n; M) \leq H(U^n) - H(M|X^n) - n \min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) + n\delta + \zeta n \min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s).$$

We now use

$$\begin{aligned} I(X^n; U^n) &= H(X_s^n) - H(X^n|U^n) \\ &\leq H(X_s^n) - H(X^n|U^nT) \\ &\leq H(X_s^n) - H(X^n|U^nT = 1)(1 - \epsilon') \\ &= H(X_s^n) - H(X_s^n|U_{\hat{s}}^nT = 1)(1 - \epsilon') \\ &= H(X_s^n) - H(X_s^n|U_{\hat{s}}^nT = 1)(1 - \epsilon') - H(X_s^n|U_{\hat{s}}^nT = 0)\epsilon' + H(X_s^n|U_{\hat{s}}^nT = 0)\epsilon' \\ &\leq I(X_s^n; U_{\hat{s}}^nT) + \epsilon' \log|\mathcal{X}|n \\ &= \epsilon' \log|\mathcal{X}|n + I(X_s^n; U_{\hat{s}}^n) + I(T; X_s^n|U_{\hat{s}}^n) \\ &\leq \epsilon' \log|\mathcal{X}|n + I(X_s^n; U_{\hat{s}}^n) + \log 2, \end{aligned}$$

where $T = t(X^n)$, $t\colon \mathcal{X}^n \to \{0, 1\}$

$$t(x^n) = \begin{cases} 1, & x^n \in \mathcal{T}^n_{P_{X_s},\xi'}, \\ 0, & \text{else.} \end{cases}$$

Thus, ϵ' is arbitrarily small for large n.

Thus, we get

$$\begin{aligned} I(X^n; M) \leq\ & H(U^n) - H(M|X^n) - n \inf_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) \\ & + n\delta + \zeta n \inf_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) + \epsilon' \log|\mathcal{X}|n + I(X_s^n; U_{\hat{s}}^n) + \log 2 - I(X^n; U^n). \end{aligned} \tag{11}$$

Again, using ([3], Problem 3.1), we get

$$\begin{aligned} H(U^n) - H(M|X^n) - I(X^n; U^n) &= H(U^n|X^n) - H(M|X^n) \\ &= H(U^nM|X^n) - H(M|X^n) \\ &= H(U^n|MX^n). \end{aligned}$$

We also know that

$$\begin{aligned} 0 &\leq I(U^n;Y^n|X^nM) \\ &= H(Y^n|X^nM) - H(Y^n|X^nU^nM) \\ &= H(Y^n|X^nM) - H(Y^n|X^nU^n) \\ &\leq H(Y^n|X^n) - H(Y^n|X^nU^n) \\ &= I(Y^n;U^n|X^n) = 0. \end{aligned}$$

Here, we use ([3], Problem 3.1) and $M - X^n - Y^n$. Thus,

$$I(U^n;X^nY^nM) = I(U^n;X^nM) = I(U^n;Y^nM) + I(U^n;X^n|Y^nM).$$

Thus,

$$I(U^n;X^nM) \geq I(U^n;Y^nM).$$

It follows that

$$H(U^n|MX^n) \leq H(U^n|MY^n). \tag{12}$$

Now, we bound the right hand side of Inequality (11) using Inequality (12) and use Fano's inequality. Thus, we have

$$\begin{aligned} &\tfrac{1}{n}I(X^n;M) \leq \sup_{s\in\mathcal{I}(\hat{s})} I(X_s;U_{\hat{s}}) - I(U_{\hat{s}};Y_s) \\ &+ \delta + \zeta I(U_{\hat{s}};Y_s) + \epsilon' \log|\mathcal{X}| + \tfrac{1}{n}\log 2 + P_e \log(|\mathcal{U}|-1) + \tfrac{h(P_e)}{n}. \end{aligned} \tag{13}$$

Here, we use

$$I(X_s;U_{\hat{s}}) - \inf_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) = \sup_{s\in\mathcal{I}(\hat{s})} I(X_s;U_{\hat{s}}) - I(U_{\hat{s}};Y_s)$$

as $I(X_s;U_{\hat{s}})$ is constant for all $s \in \mathcal{I}(\hat{s})$.

Using these results, we conclude from Inequalities (10) and (13) that

$$\mathcal{R}^{(PSCA)}(\mathfrak{S}) \supseteq \bigcup_{U_{\hat{s}_1},\cdots,U_{\hat{s}_{|\hat{\mathcal{S}}|}}} \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \mathcal{R}^{(PSCA)}_{\hat{s}}(\mathfrak{S},U_{\hat{s}}).$$

Using the distributive law for sets, we can see that this is equivalent to

$$\mathcal{R}^{(PSCA)}(\mathfrak{S}) \supseteq \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \mathcal{R}^{(PSCA)}_{\hat{s}}(\mathfrak{S},U_{\hat{s}})$$

(see Appendix B). We now consider the converse. Assume X^nY^n are distributed i.i.d. according to $P_{X_sY_s}$ for an arbitrary $s \in \mathcal{S}$. The following calculations hold for all $s \in \mathcal{S}$. Similarly to the converse part of the proof of ([7], Theorem 3.1), we have

$$\begin{aligned} \log|\mathcal{K}| &\overset{(a)}{=} H(K) = I(K;\hat{K}) + H(K|\hat{K}) \\ &\overset{(b)}{\leq} I(K;MY^n) + F = I(K;M) + I(K;Y^n|M) + F \\ &\overset{(c)}{\leq} I(Y^n;MK) + F = \sum_{i=1}^{n} I(KMY^{i-1};Y_i) + F, \end{aligned}$$

where we use Equation (6) for (a), Fano's inequality with $F = \delta n \log|\mathcal{K}| + 1$ and the data processing inequality in combination with $K - MY^n - \hat{K}$, which follows from the definition of the compound authentication protocol for (b) and Equation (7) for (c). From the definition of the compound authentication protocol, we also know that $Y^n - X^n - MK$. Using the definition of Markov chains, this implies $Y^{i-1} - X^{i-1} - MKY_i$ for all $i \in \{2 \cdots n\}$ (see Appendix C). (From $Y^n - X^{i-1}X_i^n - MK$, we get $Y^{i-1}Y_i - X^{i-1} - MKX_i^n$ using Implications (A11) and (A13). Then, we use Implication (A12) to get $Y^{i-1} - X^{i-1}Y_i - MK$ and from this we get the desired result using Implication (A13).)

The equation

$$I(Y_iKM; X^{i-1}Y^{i-1}) = I(Y_iKM; X^{i-1})$$

is equivalent to $Y^{i-1} - X^{i-1} - MKY_i$ ([3], Definition 3.9). This is equivalent to

$$H(Y_i|KMX^{i-1}Y^{i-1}) = H(Y_i|KMX^{i-1}) + (H(KM|X^{i-1}) - H(KM|X^{i-1}Y^{i-1})).$$

Thus, $H(Y_i|KMY^{i-1}) \geq H(Y_i|KMX^{i-1})$. Thus, we have

$$I(KMY^{i-1}; Y_i) \leq I(KMX^{i-1}; Y_i), \tag{14}$$

so

$$\log|\mathcal{K}| \leq \sum_{i=1}^{n} I(KMX^{i-1}; Y_i) + F.$$

Now, we define $U_i = KMX^{i-1}$ for all $i \in \{1 \cdots n\}$. This implies $U_i - X_i - Y_i$ for all $i \in \{1 \cdots n\}$, which can again be seen using the results from Appendix C. Let Q be a time sharing RV independent of all others and uniformly distributed on $\mathcal{Q} = \{1 \cdots n\}$ and let $U = QU_Q$, $X = X_Q$ and $Y = Y_Q$. Then,

$$P_{UXY}((u,q),x,y) = P_{QU_qX_qY_q}(q,u,x,y) \overset{(a)}{=} P_{QU_q|X_q}(u,q|x)P_{X_qY_q}(x,y)$$

for all $(u,q,x,y) \in \mathcal{U}_q \times \mathcal{Q} \times \mathcal{X} \times \mathcal{Y}$, where (a) follows from $U_q - X_q - Y_q$ and the independence of Q. We have

$$P_{XY}(x,y) = \sum_{q,u} P_{QU_qX_qY_q}(q,u,x,y) = \sum_{i=1}^{n} \tfrac{1}{n} P_{X_iY_i}(x,y) \overset{(a)}{=} P_{X_sY_s}(x,y) = P_{X_qY_q}(x,y) \tag{15}$$

for an arbitrary $q \in \mathcal{Q}$ and $(x,y) \in \mathcal{X} \times \mathcal{Y}$, where (a) follows as $P_{X_iY_i} = P_{X_sY_s}$ for all $i \in \mathcal{Q}$ as the RVs X^nY^n are generated i.i.d. We also have for all $(u,q,x) \in \mathcal{U}_q \times \mathcal{Q} \times \mathcal{X}$

$$P_{U|X}(u,q|x) = \frac{\sum_{y\in\mathcal{Y}} P_{QU_qX_qY_q}(q,u,x,y)}{P_X(x)} = \frac{P_{QU_qX_q}(q,u,x)}{P_{X_q}(x)} = P_{QU_q|X_q}(q,u|x).$$

Thus, $P_{UXY}((u,q),x,y) = P_{XY}(x,y)P_{U|X}(u,q|x)$, which means $U - X - Y$. We also have

$$\begin{aligned}
\log|\mathcal{K}| &\leq \sum_{i=1}^{n} I(U_i,Y_i) + F = n\sum_{i=1}^{n} \tfrac{1}{n} I(U_Q,Y|Q=i) + F \\
&= nI(U_Q;Y|Q) + F = nH(Y|Q) - H(Y|U_QQ) + F \\
&\leq n(H(Y) - H(Y|U_QQ)) + F = nI(U_QQ;Y) + F = nI(U;Y) + F.
\end{aligned}$$

Thus, using the definition of F, we get

$$\tfrac{1}{n}\log|\mathcal{K}| \leq (1-\delta)^{-1}(I(U;Y) + \tfrac{1}{n}),$$

which implies

$$\frac{1}{n}\log|\mathcal{K}| \leq I(U;Y) + \delta \tag{16}$$

for $\delta > 0$ and n large enough. We also consider

$$\begin{aligned} I(X^n;M) &= H(M) - H(M|X^n) \\ &\geq H(M|Y^n) - H(KM|X^n) \\ &= H(KM|Y^n) - H(K|Y^nM) - H(KM|X^n). \end{aligned}$$

From the definition of the compound storage model, we know $K - MY^n - \hat{K}$. Using the data processing inequality, we get $I(K;MY^n) \geq I(K;\hat{K})$, which means $H(K|MY^n) \leq H(K|\hat{K}) \leq F$, where the last inequality follows from Fano's inequality. Thus,

$$\begin{aligned} I(X^n;M) &\geq H(KM|Y^n) - H(KM|X^n) - F \\ &= I(KM;X^n) - I(KM;Y^n) - F \\ &= \sum_{i=1}^{n} I(KM;X_i|X^{i-1}) - \sum_{i=1}^{n} I(KM;Y_i|Y^{i-1}) - F \\ &\overset{(a)}{=} \sum_{i=1}^{n} I(KMX^{i-1};X_i) - \sum_{i=1}^{n_k} I(KMY^{i-1};Y_i) - F \\ &\overset{(b)}{\geq} \sum_{i=1}^{n} I(KMX^{i-1};X_i) - \sum_{i=1}^{n} I(KMX^{i-1};Y_i) - F, \end{aligned}$$

where (a) follows as X_i and Y_i are i.i.d. and (b) follows from Inequality (14). With our definition of U, X and Y and the same argumentation as before, we get

$$\begin{aligned} \frac{1}{n}I(X^n;M) &\geq I(U;X) - I(U;Y) - \frac{F}{n} \\ &\overset{(a)}{\geq} I(U;X) - I(U;Y) - \delta \end{aligned} \tag{17}$$

for n large enough, where, for (a), we use the definition of F and Inequality (16). We have for all $(u,q,x) \in \mathcal{U}_q \times \mathcal{Q} \times \mathcal{X}$

$$\begin{aligned} P_{UX}((q,u),x) &= P_Q(q)P_{U_qX_q}(u,x) \\ &= P_{KMX^{q-1}X_q}(k,m,x^{q-1},x_q)P_Q(q) \\ &= P_Q(q)\sum_{x_{q+1}^n} P_{KMX^n}(k,m,x^n) \\ &\overset{(a)}{=} P_Q(q)\sum_{x_{q+1}^n} P_{X^n}(x^n)P_{M|X^n}(m|x^n)P_{K|X^n}(k|x^n) \\ &= P_Q(q)\sum_{x_{q+1}^n} P_{X^n}(x^n)\Theta(x^n)\Phi(x^n), \end{aligned} \tag{18}$$

where (a) follows from $M - X^n - K$, which follows from the definition of the compound authentication protocol. As P_{X^n} is the same for all $s \in \mathcal{I}(\hat{s})$, $\hat{s} \in \hat{\mathcal{S}}$, this result implies that P_{UX} is the same for all $s \in \mathcal{I}(\hat{s})$, $\hat{s} \in \hat{\mathcal{S}}$. We get the bounds (16) and (17) for each $s \in \mathcal{S}$. We denote the corresponding RVs UXY by $U_sX_sY_s$ for all $s \in \mathcal{S}$. The joint distribution of X_sY_s is $P_{X_sY_s} \in \mathfrak{S}$ as we see from Equation (15). Thus, Equation (18) and the Inequalities (16) and (17) for all $s \in \mathcal{S}$ imply

$$\mathcal{R}_{PSCS}(\mathfrak{S}) \subseteq \bigcup_{U_{\hat{s}_1},\cdots,U_{\hat{s}_{|\hat{\mathcal{S}}|}}} \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \mathcal{R}_{\hat{s}}^{(PSCA)}(\mathfrak{S},U_{\hat{s}}).$$

We again use the distributive law for sets to get our result. The bounds on the cardinality of the alphabet of the auxiliary random variables can be derived as in [19]. □

Remark 9. *This result implies Theorem 2 as we use a deterministic decoder for the achievability proof.*

Remark 10. *In [19], the authors also derive the compound capacity region for $|\mathcal{S}| < \infty$, but, in contrast to this work, they consider deterministic protocols and require strong secrecy instead of perfect secrecy when defining achievability. This compound capacity region equals $\mathcal{R}_{PSCA}(\mathfrak{S})$.*

8. Secure Storage

We now discuss some other applications of the already proven results apart from authentication. For this purpose, we take a look at some results for secure storage from [13,14], which follow directly from our results for authentication. Here, we again consider compound sets $\mathfrak{S}$ with $|\hat{\mathcal{S}}| < \infty$.

In [13], we consider the following model for secure storage with source uncertainty, where the corresponding scenario is depicted in Figure 3.

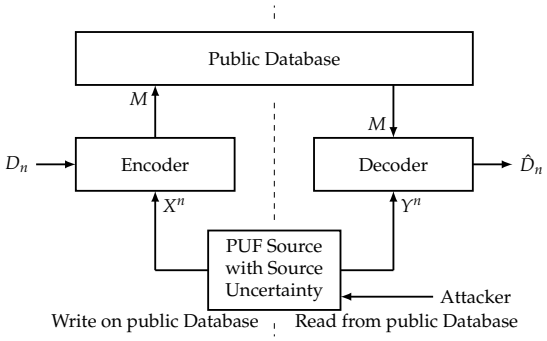

Figure 3. Secure storage process with source uncertainty (as considered in [13]).

Definition 10. *Let $n \in \mathbb{N}$. The* compound storage model *consists of a set $\mathfrak{S} \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ of DMMSs with generic variables X_sY_s, $s \in \mathcal{S}$, (all on the same alphabets $\mathcal{X}$ and $\mathcal{Y}$), a source $P_{D_n} \in \mathcal{P}(\mathcal{D}_n)$ that puts out a RV D_n, the (possibly randomized) encoder $\Phi_n \colon \mathcal{X}^n \times \mathcal{D}_n \to \mathcal{M}$ and the (possibly randomized) decoder $\Psi_n \colon \mathcal{Y}^n \times \mathcal{M} \to \hat{\mathcal{D}}_n$. Let X^n and Y^n be the output of one of the DMMSs in $\mathfrak{S}$, i.e., $P_{XY} = P_{X_sY_s}$ for an $s \in \mathcal{S}$, but s is not known. D_n is independent of X^nY^n. The RV M is generated from X^n and D_n using Φ_n. The RV $\hat{D}_n$ is generated from Y^n and M using Ψ_n. We use the term* compound storage protocol *for (Φ_n, Ψ_n). Additionally, it holds that, for all $\delta > 0$, there is an $n_0 = n_0(\delta)$ such that for all $n \geq n_0$*

$$\tfrac{1}{n} D(P_{D_n} \| U_{\mathcal{D}_n}) < \delta.$$

We define achievability for this model.

Definition 11. *A tuple (R, L), $R, L \geq 0$, is an* achievable *storage rate versus privacy-leakage rate pair for the compound storage model if for every $\delta > 0$ there is an $n_0 = n_0(\delta)$ such that for all $n \geq n_0$ there exists a compound storage protocol such that for all $s \in \mathcal{S}$*

$$\begin{aligned} Pr(D_n = \hat{D}_n) &\geq 1 - \delta, \\ I(M; D_n) &= 0, \\ \tfrac{1}{n} \log |\mathcal{D}_n| &\geq R - \delta, \\ \tfrac{1}{n} I(M; X^n) &\leq L + \delta, \end{aligned}$$

where $P_{XY} = P_{X_sY_s}$. We denote the corresponding storage protocols by PSCS-Protocols (Perfect-Secrecy-Compound-Storage-Protocols).

Definition 12. *The set of achievable rate pairs that are achievable using PSCS-Protocols is called the* compound capacity region *$\mathcal{R}_{PSCS}(\mathfrak{S})$.*

We then can prove the following result.

Theorem 6. *It holds that*

$$\mathcal{R}_{PSCS}(\mathfrak{S}) = \mathcal{R}_{PSCA}(\mathfrak{S}).$$

Remark 11. *The compound storage model is essentially equivalent to a compound version of the chosen secret system in [7]. For this reason, Theorem 6 follows using the same approach as the authors of [7].*

We combine source compression and secure storage in [14] by considering the following model, which models the scenario depicted in Figure 4.

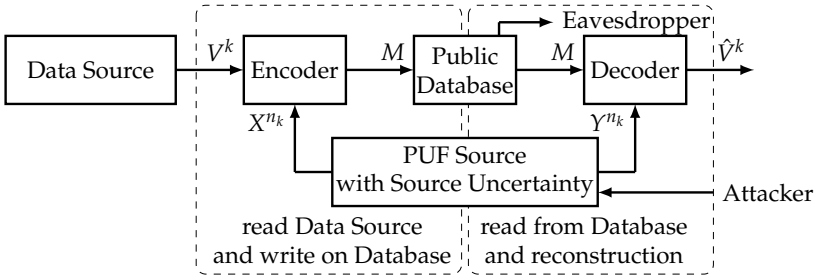

Figure 4. Secure storage of a source (as considered in [14]).

Definition 13. *Let $k, n_k \in \mathbb{N}$. The* compound source storage model *consists of a set $\mathfrak{S} \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ of DMMSs with generic variables X_sY_s, $s \in \mathcal{S}$, (all on the same alphabets $\mathcal{X}$ and $\mathcal{Y}$), a general source $\mathbf{V}$ [20] that fulfills the strong converse property, the (possibly randomized) encoder $\Phi_k \colon \mathcal{X}^{n_k} \times \mathcal{V}^k \to \mathcal{M}$ and the (possibly randomized) decoder $\Psi_k \colon \mathcal{Y}^{n_k} \times \mathcal{M} \to \hat{\mathcal{V}}^k$. Let X^{n_k} and Y^{n_k} be the output of one of the DMMSs in $\mathfrak{S}$, i.e., $P_{XY} = P_{X_sY_s}$ for an $s \in \mathcal{S}$, but s is not known. The RV M is generated from X^{n_k} and V^k using Φ_k. The RV $\hat{V}^k$ is generated from Y^{n_k} and M using Ψ_k. We use the term* compound source storage protocol *for (Φ_k, Ψ_k).*

For this model, we define achievability where we consider the output of the PUF source as a resource.

Definition 14. *A tuple (B, L), $B, L \geq 0$, is an* achievable *performance pair for the compound source storage model if, for every $\delta > 0$, there is a $k_0 = k_0(\delta)$ such that, for all $k \geq k_0$, there exists a compound source storage protocol such that, for all $s \in \mathcal{S}$,*

$$\begin{aligned} &Pr(V^k = \hat{V}^k) \geq 1 - \delta, \\ &I(M; V^k) = 0, \\ &\tfrac{n_k}{k} \leq B + \delta, \\ &\tfrac{1}{n_k} I(M; X^{n_k}) \leq L + \delta, \end{aligned}$$

where $P_{XY} = P_{X_sY_s}$. We denote the corresponding compound source storage protocols by PSCSS-Protocols (Perfect-Secrecy-Compound-Source-Storage-Protocols).

Definition 15. *The set of achievable performance pairs that are achievable using PSCSS-Protocols is called the* optimal performance region $\mathcal{R}_{PSCSS}(\mathfrak{S}, \mathbf{V})$.

We then can prove the following results.

Theorem 7. *It holds that*

$$\mathcal{R}_{PSCSS}(\mathfrak{S},\mathbf{V}) \supseteq \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \{(B,L)\colon B \geq \tfrac{\bar{H}(\mathbf{V})}{\inf_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s)}, L \geq \sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_s) - I(U_{\hat{s}};Y_s)\}$$
$$\overset{(a)}{=} \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \mathcal{R}_{\hat{s}}^{(PSCSS)}(\mathfrak{S},\mathbf{V},U_{\hat{s}}),$$

where for (a) *we define* $\mathcal{R}_{\hat{s}}^{(PSCSS)}(\mathfrak{S},\mathbf{V},U_{\hat{s}})$ *appropriately. For all* $\hat{s}\in\hat{\mathcal{S}}$*, the union is over all RVs* $U_{\hat{s}}$ *such that, for all* $s\in\mathcal{I}(\hat{s})$*, we have* $U_{\hat{s}} - X_s - Y_s$*.*

Theorem 8. *For stationary ergodic sources* $\mathbf{V}$*, it holds that*

$$\mathcal{R}_{PSCSS}(\mathfrak{S},\mathbf{V}) = \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \mathcal{R}_{\hat{s}}^{(PSCSS)}(\mathfrak{S},\mathbf{V},U_{\hat{s}}).$$

For all $\hat{s}\in\hat{\mathcal{S}}$*, the union is over all RVs* $U_{\hat{s}}$ *such that, for all* $s\in\mathcal{I}(\hat{s})$*, we have* $U_{\hat{s}} - X_s - Y_s$*. For* $|\mathcal{S}|<\infty$*, we only have to consider RVs* $U_{\hat{s}}$ *with* $|\mathcal{U}_{\hat{s}}| \leq |\mathcal{X}| + |\mathcal{I}(\hat{s})|$*.*

9. Conclusions

We derived the capacity region for the (compound) authentication model requiring perfect secrecy and uniform distribution of the key generated for authentication and compared the result to existing results where only strong secrecy and a weaker condition on the key distribution is required. The two capacity regions are the same. We could prove this result by allowing for randomized encoders, which are not necessarily used when deriving the capacity region corresponding to the weaker definition of achievability. We saw that we can use the results for authentication to prove corresponding results for secure storage.

As already mentioned, compound sources do not only model source uncertainty but also model attacks where an attacker can influence parameters of the source while the legitimate parties do not know which parameters the attacker chose. It is essential that in this scenario the parameter is constant for all symbols read from the source. An attack where the parameter can be varied while the source is used is fundamentally stronger. A characterization of achievable rates for this attack scenario is not known, except for the source model for secret key generation, which has been derived in [21]. For an overview of these types of attacks, see [22]. Recently, the corresponding problem for wiretap channels could be solved [23,24]. For the source model, the attacker can choose his strategy depending on the public data, which is a difficulty that does not appear for wiretap channels. Nevertheless the authors hope that, using techniques from the works concerning the wiretap channel, the open problem for the source model can be solved.

Acknowledgments: Funding is acknowledged from the German Research Foundation (DFG) via grant BO 1734/20-1 and from the Federal Ministry of Education and Research (BMBF) via grant 16KIS0118K. Holger Boche would like to thank Rainer Plaga, Federal Office for Information Security (BSI), for the discussion on PUFs and issues concerning different secrecy measures.

Author Contributions: Sebastian Baur and Holger Boche conceived this study and derived the results. Sebastian Baur wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Theorem 4

Proof. We prove the result for compound codes with the additional constraint on the decoding sets that, for $\zeta>0$, it holds that

$$\phi^{-1}(m) \subset \bigcup_{W\in\mathcal{W}} \mathcal{T}_{W,\zeta}^{n}(f(m)) \tag{A1}$$

for all messages $m \in \mathcal{M}_f$. Additionally, for $\zeta' > 0$, we require

$$f(m) \in \tilde{A} = A \cap \mathcal{T}^n_{P,\zeta'} \tag{A2}$$

for all $m \in \mathcal{M}_f$. First, consider the case that $\mathcal{W}$ is a finite set. Let (f, ϕ) be such a code that can not be extended. Thus, for all $x^n \in \tilde{A}$, there is a $W \in \mathcal{W}$ such that

$$W^n(\bigcup_{\tilde{W}\in\mathcal{W}} \mathcal{T}^n_{\tilde{W},\zeta}(x^n) \setminus B | x^n) < 1 - \epsilon, \tag{A3}$$

where $B = \bigcup_{m\in\mathcal{M}_f} \phi^{-1}(m)$. It also holds that

$$P^n(\tilde{A}) \geq P^n(A) + P^n(\mathcal{T}^n_{P,\zeta'}) - 1 \geq \eta/2$$

for n large enough. We now consider the set

$$\tilde{A}_W = \{x^n \in \tilde{A} \colon W^n(\bigcup_{\tilde{W}\in\mathcal{W}} \mathcal{T}^n_{\tilde{W},\zeta}(x^n) \setminus B | x^n) < 1 - \epsilon\}.$$

We know $\bigcup_{W\in\mathcal{W}} \tilde{A}_W = \tilde{A}$, as for all $x^n \in \tilde{A}$ there is at least one $W \in \mathcal{W}$ with Inequality (A3). Thus,

$$\eta/2 \leq P^n(\bigcup_{W\in\mathcal{W}} \tilde{A}_W) \leq \sum_{W\in\mathcal{W}} P^n(\tilde{A}_W) \leq |\mathcal{W}| \max_{W\in\mathcal{W}} P^n(\tilde{A}_W).$$

Thus, there is a $\bar{W} \in \mathcal{W}$ such that for all $x^n \in \tilde{A}_{\bar{W}}$

$$\bar{W}^n(\bigcup_{\tilde{W}\in\mathcal{W}} \mathcal{T}^n_{\tilde{W},\zeta}(x^n) \setminus B | x^n) < 1 - \epsilon$$

and

$$P^n(\tilde{A}_{\bar{W}}) \geq \tfrac{\eta}{2|\mathcal{W}|}.$$

Thus,

$$\begin{aligned}&\bar{W}^n(B^c|x^n) + \bar{W}^n(\bigcup_{\tilde{W}\in\mathcal{W}} \mathcal{T}^n_{\tilde{W},\zeta}(x^n)|x^n) - \bar{W}^n(B^c \cup \bigcup_{\tilde{W}\in\mathcal{W}} \mathcal{T}^n_{\tilde{W},\zeta}(x^n)|x^n)\\ &= \bar{W}^n(\bigcup_{\tilde{W}\in\mathcal{W}} \mathcal{T}^n_{\tilde{W},\zeta}(x^n) \setminus B | x^n) < 1 - \epsilon,\end{aligned}$$

which means

$$\bar{W}^n(B|x^n) > \epsilon - \delta$$

for all $x^n \in \tilde{A}_{\bar{W}}$ as $\bar{W}^n(B^c|x^n) = 1 - \bar{W}^n(B|x^n)$, $\bar{W}^n(\bigcup_{\bar{W}\in\mathcal{W}} \mathcal{T}^n_{\bar{W},\zeta}(x^n)|x^n) \geq 1 - \delta$ for $\delta > 0$ and n large enough and $\bar{W}^n(B^c \cup \bigcup_{\bar{W}\in\mathcal{W}} \mathcal{T}^n_{\bar{W},\zeta}(x^n)|x^n) \leq 1$. Thus, we have

$$\begin{aligned}\bar{W}^n(B \cap \bigcup_{\bar{x}^n\in\mathcal{T}^n_{P,\zeta'}} T^n_{\bar{W},\zeta}(\bar{x}^n)|x^n) &\geq \bar{W}^n(B|x^n) + \bar{W}^n(\bigcup_{\bar{x}^n\in\mathcal{T}^n_{P,\zeta'}} T^n_{\bar{W},\zeta}(\bar{x}^n)|x^n) - 1\\ &\geq \epsilon - \delta + (1 - \xi) - 1 = \epsilon - \xi - \delta\end{aligned}$$

for all $x^n \in \tilde{A}_{\bar{W}}$, $\xi > 0$ and n large enough. (We choose ϵ, δ and ξ such that $\epsilon - \xi - \delta > 0$.) Thus, $B' = B \cap \bigcup_{\bar{x}^n \in \mathcal{T}^n_{P,\xi'}} T^n_{\bar{W},\xi}(\bar{x}^n)$ is an $\epsilon - \xi - \delta$ image of $\tilde{A}_{\bar{W}}$ (see [3]). Thus,

$$|B \cap \bigcup_{\bar{x}^n \in \mathcal{T}^n_{P,\xi'}} T^n_{\bar{W},\xi}(\bar{x}^n)| \geq g_{\bar{W}^n}(\tilde{A}_{\bar{W}}, \epsilon - \xi - \delta),$$

where $g_{\bar{W}^n}(\tilde{A}_{\bar{W}}, \epsilon - \xi - \delta)$ is defined as in [3]. We have

$$\begin{aligned}
(P\bar{W})^n(B') &= \sum_{y^n \in B'} \prod_{i=1}^{n} \sum_{a \in \mathcal{X}} P(a)\bar{W}(y_i|a) \\
&\overset{(a)}{=} \sum_{y^n \in B'} \sum_{x^n \in \mathcal{X}^n} \prod_{i=1}^{n} P(x_i)\bar{W}(y_i|x_i) \\
&\geq \sum_{y^n \in B'} \sum_{x^n \in \tilde{A}_{\bar{W}}} P^n(x^n)\bar{W}^n(y^n|x^n) \\
&= \sum_{x^n \in \tilde{A}_{\bar{W}}} P^n(x^n) \sum_{y^n \in B'} \bar{W}^n(y^n|x^n) \\
&\geq (\epsilon - \delta - \xi)P^n(\tilde{A}_{\bar{W}}) \geq \eta/2(\epsilon - \delta - \xi)\tfrac{1}{|\mathcal{W}|},
\end{aligned}$$

where (a) can be shown with induction. Using ([3], Lemma 2.14), we get for n large enough

$$\tfrac{1}{n}\log|B'| \geq H(P\bar{W}) - (\gamma + \tfrac{1}{n}\log|\mathcal{W}|) \tag{A4}$$

with $\gamma > 0$. Additionally, we have

$$\begin{aligned}
|B'| &\overset{(a)}{=} |\bigcup_{m \in \mathcal{M}_f} \phi^{-1}(m) \cap \bigcup_{x^n \in \mathcal{T}^n_{P,\xi'}} \mathcal{T}^n_{\bar{W},\xi}(x^n)| \\
&= |\bigcup_{m \in \mathcal{M}_f} \left[\phi^{-1}(m) \cap \bigcup_{x^n \in \mathcal{T}^n_{P,\xi'}} \mathcal{T}^n_{\bar{W},\xi}(x^n)\right]| \\
&\leq \sum_{m \in \mathcal{M}_f} |\phi^{-1}(m) \cap \bigcup_{x^n \in \mathcal{T}^n_{P,\xi'}} \mathcal{T}^n_{\bar{W},\xi}(x^n)| \\
&\overset{(b)}{\leq} \sum_{m \in \mathcal{M}_f} |\bigcup_{W \in \mathcal{W}} \mathcal{T}^n_{W,\zeta}(f(m)) \cap \bigcup_{x^n \in \mathcal{T}^n_{P,\xi'}} \mathcal{T}^n_{\bar{W},\zeta}(x^n)|,
\end{aligned}$$

where (a) follows from the definition of B and (b) follows from Subset Relationship (A1). We now define

$$\mathcal{W}^*_m = \{W \in \mathcal{W} \colon \mathcal{T}^n_{W,\zeta}(f(m)) \cap \bigcup_{x^n \in \mathcal{T}^n_{P,\xi'}} \mathcal{T}^n_{\bar{W},\zeta}(x^n) \neq \emptyset\}.$$

As

$$\mathcal{T}^n_{\bar{W},\xi}(f(m)) \cap \bigcup_{x^n \in \mathcal{T}^n_{P,\xi'}} \mathcal{T}^n_{\bar{W},\zeta}(x^n) \neq \emptyset$$

for all $m \in \mathcal{M}_f$, which follows form Relation (A2), we have

$$|B'| \leq \sum_{m \in \mathcal{M}_f} \max_{W \in \mathcal{W}^*_m} |\mathcal{T}^n_{W,\zeta}(f(m))| \cdot |\mathcal{W}|.$$

Let

$$W^* = \arg \max_{W \in \bigcup_{m \in \mathcal{M}_f} \mathcal{W}_m^*} |\mathcal{T}_{W,\zeta}^n(f(m))|.$$

Thus, we get the upper bound

$$|B'| \leq |\mathcal{M}_f| \exp(n(H(W^*|P) + \gamma' + \tfrac{\log|\mathcal{W}|}{n})), \tag{A5}$$

$\gamma' > 0$ ([3], Lemma 2.13).

For all $W \in \mathcal{W}_m^*$ and all $m \in \mathcal{M}_f$ there is a $y^n \in \mathcal{Y}^n$ such that $y^n \in \mathcal{T}_{W,\zeta}^n(f(m))$ and $y^n \in \mathcal{T}_{\bar{W},\zeta}^n(x^n)$ for a $x^n \in \mathcal{T}_{P,\zeta'}^n$. Using Relation (A2), we have $y^n \in \mathcal{T}_{PW,(\zeta+\zeta')|\mathcal{X}|}^n$ and $y^n \in \mathcal{T}_{P\bar{W},(\zeta+\zeta')|\mathcal{X}|}^n$ (see ([3], Lemma 2.10)). Let $\zeta'' = (\zeta + \zeta')|\mathcal{X}|$. Thus,

$$\begin{aligned}
\|PW - P\bar{W}\|_1 &= \sum_{b \in \mathcal{Y}} |PW(b) - P\bar{W}(b)| \\
&= \sum_{b \in \mathcal{Y}} |PW(b) - N(b|y^n)/n + N(b|y^n)/n - P\bar{W}(b)| \\
&\leq \sum_{b \in \mathcal{Y}} |PW(b) - N(b|y^n)/n| + |N(b|y^n)/n - P\bar{W}(b)| \leq 2|\mathcal{Y}|\zeta''.
\end{aligned}$$

Using ([3], Lemma 2.7), we have $|H(PW) - H(P\bar{W})| \leq 2|\mathcal{Y}|\zeta'' \log \frac{1}{2\zeta''}$ for all $W \in \mathcal{W}_m^*$ and all $m \in \mathcal{M}_f$. Using Inequalities (A4), (A5) and the fact that $W^* \in \mathcal{W}_m^*$ for a $m \in \mathcal{M}_f$, we get for γ, γ', ζ and ζ' small enough and n large enough

$$\begin{aligned}
\tfrac{1}{n} \log |\mathcal{M}_f| &\geq H(P\bar{W}) - H(W^*|P) - \gamma - \gamma' - 2\tfrac{\log|\mathcal{W}|}{n} \\
&\geq H(PW^*) - 2|\mathcal{Y}|\zeta'' \log \tfrac{1}{2\zeta''} - H(W^*|P) - \gamma - \gamma' - 2\tfrac{\log|\mathcal{W}|}{n} \\
&\geq I(P; W^*) - \tau \geq \min_{W \in \mathcal{W}} I(P; W) - \tau.
\end{aligned} \tag{A6}$$

Now, consider the case of an infinite set $\mathcal{W}$. Let $M \in \mathbb{N}$, $M \geq 2|\mathcal{Y}|^2$. We construct the set $\mathcal{W}^*$ of channels $W^* \colon \mathcal{X} \to \mathcal{Y}$ with the following properties. For all $W \in \mathcal{W}$, there is a $W^* \in \mathcal{W}^*$ with

$$|W(y|x) - W^*(y|x)| \leq \tfrac{|\mathcal{Y}|}{M} \tag{A7}$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$W(y|x) \leq W^*(y|x) e^{2|\mathcal{Y}|^2/M} \tag{A8}$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and

$$|\mathcal{W}^*| \leq (1 + M)^{|\mathcal{X}||\mathcal{Y}|}. \tag{A9}$$

Such a construction is possible as described in [18]. Using Inequalities (A9) and (A6), we know that there is a compound (n, ϵ')-code, $\epsilon > \epsilon' > 0$, for $\mathcal{W}^*$ with

$$\tfrac{1}{n} \log |\mathcal{M}_f| \geq \min_{W \in \mathcal{W}^*} I(P; W) - \tau$$

if M depends on n polynomially. We now show that this code is a compound (n, ϵ)-code for $\mathcal{W}$ with

$$\tfrac{1}{n} \log |\mathcal{M}_f| \geq \inf_{W \in \mathcal{W}} I(P; W) - \tau.$$

Let $W^* = \arg\min_{W \in \mathcal{W}^*} I(P;W)$ and let $W \in \mathcal{W}$ be the W corresponding to W^*. Then, we have

$$\inf_{W \in \mathcal{W}} I(P;W) \overset{(a)}{\leq} I(P;W) \overset{(b)}{\leq} I(P;W^*) + \beta \overset{(c)}{=} \min_{W \in \mathcal{W}^*} I(P;W) + \beta,$$

$\beta > 0$, where (a) follows from the definition of the infimum, (b) follows as Inequality (A7) implies

$$\|W(\cdot|a) - W^*(\cdot|a)\|_1 \leq \tfrac{|\mathcal{Y}|^2}{M}$$

for all $a \in \mathcal{X}$. Thus, using ([3], Lemma 2.7), we have

$$\begin{aligned} |I(P;W) - I(P;W^*)| &= |H(W|P) - H(W^*|P)| \\ &= |\sum_{a \in \mathcal{X}} P(a)(H(W(\cdot|a)) - H(W^*(\cdot|a)))| \\ &\leq \sum_{a \in \mathcal{X}} P(a)|H(W(\cdot|a)) - H(W^*(\cdot|a))| \leq \tfrac{|\mathcal{Y}|^2}{M} \log \tfrac{M}{|\mathcal{Y}|}. \end{aligned}$$

For $M = n^2$, we get (b) for n large enough. Finally, (c) follows from the choice of W^*. Additionally, it holds that for each $W \in \mathcal{W}$ there is a $W^* \in \mathcal{W}^*$ with

$$W^n(y^n|x^n) \leq e^{2|\mathcal{Y}|^2 n/M} (W^*)^n(y^n|x^n),$$

which follows from Inequality (A8). Thus, for all $m \in \mathcal{M}_f$, we have

$$W^n((\phi^{-1}(m))^c|f(m)) \leq (W^*)^n((\phi^{-1}(m))^c|f(m)) e^{2|\mathcal{Y}|^2 n/M} \overset{a)}{\leq} e^{2|\mathcal{Y}|^2/n} \epsilon',$$

where (a) follows from our choice of M. Thus, for n large enough and ϵ' small enough, we have

$$W^n((\phi^{-1}(m))^c|f(m)) \leq \epsilon.$$

□

Appendix B. Equivalence of Rate Regions

We have

$$\bigcup_{U_{\hat{s}_1}, \cdots, U_{\hat{s}_{|\hat{\mathcal{S}}|}}} \bigcap_{\hat{s} \in \hat{\mathcal{S}}} \mathcal{R}_{\hat{s}}^{(PSCA)}(\mathfrak{S}, U_{\hat{s}}) \overset{(a)}{=} \bigcup_{U_{\hat{s}_1}} \bigcup_{U_{\hat{s}_2}, \cdots, U_{\hat{s}_{|\hat{\mathcal{S}}|}}} \left(\bigcap_{\hat{s} \in \hat{\mathcal{S}} \setminus \{\hat{s}_1\}} \mathcal{R}_{\hat{s}}(\mathfrak{S}, U_{\hat{s}}) \cap \mathcal{R}_{\hat{s}_1}(\mathfrak{S}, U_{\hat{s}_1}) \right),$$

where we drop the $(PSCA)$ for a shorter notation in (a). We now use the distributive law for sets and get

$$\bigcup_{U_{\hat{s}_1}} \left(\bigcup_{U_{\hat{s}_2}, \cdots, U_{\hat{s}_{|\hat{\mathcal{S}}|}}} \bigcap_{\hat{s} \in \hat{\mathcal{S}} \setminus \{\hat{s}_1\}} \mathcal{R}_{\hat{s}}(\mathfrak{S}, U_{\hat{s}}) \cap \mathcal{R}_{\hat{s}_1}(\mathfrak{S}, U_{\hat{s}_1}) \right).$$

Now, we use the distributive law again and get

$$\bigcup_{U_{\hat{s}_2}, \cdots, U_{\hat{s}_{|\hat{\mathcal{S}}|}}} \bigcap_{\hat{s} \in \hat{\mathcal{S}} \setminus \{\hat{s}_1\}} \mathcal{R}_{\hat{s}}(\mathfrak{S}, U_{\hat{s}}) \cap \bigcup_{U_{\hat{s}_1}} \mathcal{R}_{\hat{s}_1}(\mathfrak{S}, U_{\hat{s}_1}).$$

Following these steps for all $\hat{s} \in \hat{\mathcal{S}}$, we get

$$\bigcap_{\hat{s} \in \hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \mathcal{R}_{\hat{s}}^{(PSCS)}(\mathfrak{S}, U_{\hat{s}}).$$

Appendix C. Modifying Markov Chains

Theorem A1. *Let A,B,C and D be jointly distributed RVs. It holds that*

$$A - B - C \Leftrightarrow C - B - A, \tag{A10}$$
$$AB - C - D \Rightarrow B - C - D, \tag{A11}$$
$$AB - C - D \Rightarrow A - BC - D, \tag{A12}$$
$$P_{ABC}(a,b,c) = P_{AB}(a,b)P_C(c) \quad \forall (a,b,c) \in \mathcal{A} \times \mathcal{B} \times \mathcal{C},$$
$$\wedge\, A - BC - D \Rightarrow A - B - CD. \tag{A13}$$

Proof. We give a proof for each of the statements.

- We have

$$\begin{aligned} P_{ABC}(a,b,c) &\overset{(a)}{=} P_{A|B}(a|b)P_{BC}(b,c) \\ &= P_{A|B}(a|b)P_{C|B}(c|b)P_B(b) = P_{AB}(a,b)P_{C|B}(c|b) \end{aligned}$$

  for all $(a,b,c) \in \mathcal{A} \times \mathcal{B} \times \mathcal{C}$. Here, (a) follows from $A - B - C$. Thus, we see that Equivalence (A10) is true.
- We have $P_{ABCD}(a,b,c,d) = P_{AB|C}(a,b|c)P_{CD}(c,d)$ for all $(a,b,c,d) \in \mathcal{A} \times \mathcal{B} \times \mathcal{C} \times \mathcal{D}$ from $AB - C - D$. Summing both sides over all $b \in \mathcal{B}$, we get Implication (A11).
- We have

$$\begin{aligned} P_{ABCD}(a,b,c,d) &\overset{(a)}{=} P_{AB|C}(a,b|c)P_{CD}(c,d) \\ &= P_{B|C}(b,c)P_{A|BC}(a|b,c)P_{CD}(c,d) \\ &\overset{(b)}{=} P_{A|BC}(a|b,c)P_{B|CD}(b|c,d)P_{CD}(c,d) \\ &= P_{A|BC}(a|b,c)P_{BCD}(b,c,d) \end{aligned}$$

  for all $(a,b,c,d) \in \mathcal{A} \times \mathcal{B} \times \mathcal{C} \times \mathcal{D}$, where (a) follows from $AB - C - D$ and (b) from Implication (A11). This means Implication (A12) is true.
- We have

$$\begin{aligned} P_{ABCD}(a,b,c,d) &\overset{(a)}{=} P_{A|BC}(a|b,c)P_{BCD}(b,c,d) \\ &= P_{A|BC}(a|b,c)P_{D|BC}(d|b,c)P_{BC}(b,c) \\ &\overset{(b)}{=} P_{AB}(a,b)P_C(c)P_{D|BC}(d|b,c) \\ &= P_{A|B}(a|b)P_B(b)P_C(c)P_{D|BC}(d|b,c) \\ &\overset{(c)}{=} P_{A|B}(a|b)P_{BC}(b,c)P_{D|BC}(d|b,c) \\ &= P_{A|B}(a|b)P_{BCD}(b,c,d) \end{aligned}$$

  for all $(a,b,c,d) \in \mathcal{A} \times \mathcal{B} \times \mathcal{C} \times \mathcal{D}$, where (a) follows from $A - BC - D$ and (b) and (c) follow as C is independent of AB. Thus, we have Implication (A13).

□

References

1. Schaefer, R.F.; Boche, H.; Khisti, A.; Poor, H.V. *Information Theoretic Security and Privacy of Information Systems*; Cambridge University Press: Cambridge, UK, 2017.
2. Shannon, C.E. Communication theory of secrecy systems. *Bell Syst. Tech. J.* **1949**, *28*, 656–715.
3. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Cambridge University Press: Cambridge, UK, 2011.

4. Wyner, A.D. The wire-tap channel. *Bell Syst. Tech. J.* **1975**, *54*, 1355–1387.
5. Bloch, M.; Barros, J. *Physical-Layer Security: From Information Theory to Security Engineering*; Cambridge University Press: Cambridge, UK, 2011.
6. Ahlswede, R.; Csiszár, I. Common randomness in information theory and cryptography. Part I: Secret sharing. *IEEE Trans. Inf. Theory* **1993**, *39*, 1121–1132.
7. Ignatenko, T.; Willems, F.M. Biometric security from an information theoretical perspective. *Found. Trends Commun. Inf. Theory* **2012**, *7*, 135–316.
8. Grigorescu, A.; Boche, H.; Schaefer, R.F. Robust PUF based authentication. In Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS), Rome, Italy, 16–19 November 2015; pp. 1–6.
9. Lai, L.; Ho, S.-W.; Poor, H.V. Privacy-security tradeoffs in biometric security systems. In Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, IL, USA, 23–26 September 2008; pp. 268–273.
10. Boche, H.; Wyrembelski, R.F. Secret key generation using compound sources-optimal key-rates and communication costs. In Proceedings of the 2013 9th International ITG Conference on Systems, Communication and Coding (SCC), München, Germany, 21–24 January 2013.
11. Grigorescu, A.; Boche, H.; Schaefer, R.F. Robust Biometric Authentication from an Information Theoretic Perspective. *Entropy* **2017**, *19*, 480.
12. Baur, S.; Boche, H. Robust authentication and data storage with perfect secrecy. In Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Atlanta, GA, USA, 1–4 May 2017; pp. 553–558.
13. Baur, S.; Boche, H. Robust Secure Storage of Data Sources with Perfect Secrecy. In Proceedings of the IEEE Workshop on Information Forensics and Security, Rennes, France, 4–7 December 2017.
14. Baur, S.; Boche, H. Storage of general data sources on a public database with security and privacy constraints. In Proceedings of the 2017 IEEE Conference on Communications and Network Security (CNS), Las Vegas, NV, USA, 9–11 October 2017; pp. 555–559.
15. Willems, F.; Ignatenko, T. Authentication based on secret-key generation. In Proceedings of the 2012 IEEE International Symposium on Information Theory Proceedings (ISIT), Cambridge, MA, USA, 1–6 July 2012; pp. 1792–1796.
16. Gallager, R. *Information Theory and Reliable Communication*; Springer: Berlin, Germany, 1968.
17. Wolfowitz, J. *Coding Theorems of Information Theory*; Springer: Berlin, Germany, 1978.
18. Blackwell, D.; Breiman, L.; Thomasian, A.J. The capacity of a class of channels. *Ann. Math. Stat.* **1959**, *30*, 1229–1241.
19. Tavangaran, N.; Baur, S.; Grigorescu, A.; Boche, H. Compound biometric authentication systems with strong secrecy. In Proceedings of the 2017 11th International ITG Conference on Systems, Communication and Coding (SCC), Hamburg, Germany, 6–9 February 2017.
20. Han, T.S. *Information-Spectrum Methods in Information Theory*; Springer Science & Business Media: New York, NY, USA, 2013; Volume 50.
21. Boche, H.; Cai, N. Common Random Secret Key Generation on Arbitrarily Varying Source. In Proceedings of the 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS2018), Hong Kong, China, 16–20 July 2018, in press.
22. Schaefer, R.F.; Boche, H.; Poor, H.V. Secure Communication Under Channel Uncertainty and Adversarial Attacks. *Proc. IEEE* **2015** *103*, 1796–1813.
23. Wiese, M.; Nötzel, J.; Boche, H. A Channel Under Simultaneous Jamming and Eavesdropping Attack—Correlated Random Coding Capacities Under Strong Secrecy Criteria. *IEEE Trans. Inf. Theory* **2016**, *62*, 3844–3862.
24. Nötzel, J.; Wiese, M.; Boche, H. The Arbitrarily Varying Wiretap Channel—Secret Randomness, Stability, and Super-Activation. *IEEE Trans. Inf. Theory* **2016**, *62*, 3504–3531.